

Profiling Human CMV-specific T cell responses reveals novel immunogenic ORFs

1 Rekha Dhanwani^{1†}, Sandeep Kumar Dhanda^{1#}, John Pham¹, Gregory P. Williams¹,
2 John Sidney¹, Alba Grifoni¹, Gaelle Picarda^{1,3}, Cecilia S. Lindestam Arlehamn¹,
3 Alessandro Sette^{1,2,*}, Chris A Benedict^{1,3,*}

4

5 ¹ Center for Infectious Disease and Vaccine Research, La Jolla Institute for Immunology,
6 La Jolla, CA 92037, USA

7 ² Department of Medicine, University of California San Diego, La Jolla, CA 92093, USA

8 ³Center for Autoimmunity and Inflammation, La Jolla Institute for Immunology, La Jolla,
9 CA 92037, USA

10 † Present Address: Division of Extramural Activities, National Institute of Allergy and
11 Infectious Diseases, National Institute of Health, Rockville, MD 20852, USA

12 # Present Address: St. Jude Children's Research Hospital, Arlington, VA 22203

13

14 * Corresponding authors

15 Alessandro Sette

16 Email: alex@lji.org

17

18 Chris A. Benedict

19 Email: benedict@lji.org

20

21

22 **Abstract**

23 Despite the prevalence and medical significance of human cytomegalovirus (HCMV)
24 infections, a systematic analysis of the targets of T cell recognition in humans that spans
25 the entire genome and includes recently described potential novel ORFs is not available.
26 Here, we screened a library of epitopes predicted to bind HLA class II that spans over
27 350 different HCMV ORFs and includes ~150 previously described and ~200 recently
28 described potential novel ORFs using an ex vivo IFN γ fluorospot assay. We identified 235
29 unique HCMV specific epitopes derived from 100 ORFs, some previously described as
30 immunodominant and others that were not previously described to be immunogenic. Of
31 those, 41 belong to the set of recently reported novel ORFs, thus providing evidence that
32 at least some of these are actually expressed *in vivo* in humans. These data reveal that
33 the breadth of the human T cell response to HCMV is much greater than previously
34 thought. The ORFs and epitopes identified will help elucidate how T cell immunity relates
35 to HCMV pathogenesis and instruct ongoing HCMV vaccine research.

36

37 **Importance**

38 To understand the crucial role of adaptive immunity in controlling cytomegalovirus
39 infection and disease, we systematically analyzed the CMV 'ORFeome' to identify new
40 CMV epitopes targeted primarily by CD4 T cells in humans. Our study identified >200
41 new T cell epitopes derived from both canonical and novel ORFs, highlighting the
42 substantial breadth of anti-CMV T cell response and providing new targets for vaccine
43 design.

44

45 **Introduction**

46 Human cytomegalovirus (HCMV, HHV-5) is a β -herpesvirus that infects the
47 majority of the world's population. Infection in healthy individuals is characterized by a
48 primary asymptomatic phase followed by the establishment of lifelong persistence/latency
49 in several cell types (1, 2). HCMV's 236 kbp double stranded DNA genome facilitates its
50 persistence and reactivation when immunity is compromised, with both viral and cellular
51 proteins controlling viral gene expression and regulating the dynamic and reversible
52 latent-lytic cycle that develops over a lifelong infection (3, 4). Although largely persistent,
53 its reactivation in immunocompromised populations, such as transplant recipients and
54 AIDS patients, causes severe disease outcomes (5-11). Congenital infection in the
55 developing fetus is also the leading infectious cause of birth defects (12-18). Moreover,
56 the available antiviral drug therapies are insufficient and often toxic in young children (19-
57 22). Consequently, HCMV is recognized as a major public health problem and
58 development of a vaccine that prevents or at least mitigates virus-induced disease is a
59 top priority (23-25) .

60 Although both humoral and cell mediated immune responses protect against
61 HCMV infection, a considerable effort has been made towards identifying HCMV targets
62 of CTL responses due to their pivotal role in controlling HCMV disease in
63 immunocompromised individuals (26-29). However, HCMV targets of CD4+ T helper
64 cells, which amplify CTL and antibody responses or may mediate direct antiviral activity
65 themselves, remain to be explored in detail. In order to develop a successful HCMV
66 vaccine, it is imperative to assess the large number of candidate viral proteins for their
67 potential to induce robust CD4+ T cell responses.

68 Previous work from Sylwester et al. extensively characterized the canonical HCMV
69 proteins that are targeted by CD4+ and CD8+ T cell responses (30), and work by many
70 other groups have identified immunodominant epitopes derived from these that include
71 the 65kDA phosphoprotein (UL83/pp65), immediate early protein 1 (UL123), tegument
72 protein pp150 (UL32), envelope glycoprotein B (UL55), viral transcription factor IE2
73 (UL122), and major capsid protein (UL86) (31-38). However, a comprehensive analysis
74 of HCMV epitope-specific T cell responses has been challenging, mainly due to the large
75 size of virus and the evolving impact that persistent infection has on the memory pool.
76 Stern-Ginossar et al. recently reported all the HCMV RNAs found to be associated with
77 ribosomes in infected fibroblasts, increasing the potential number of ORFs the virus may
78 encode by ~3 fold (39). Here, we designed a comprehensive screening approach to
79 assess potential T cell responses against 563 of these ORFs, which included both
80 previously reported and potentially novel HCMV proteins. 2593 15-mer peptides were
81 predicted using computational algorithms, and a high throughput screen was performed
82 using an IFN γ fluorospot assay to identify epitopes targeted by both CD8+ and CD4+ T
83 cells in healthy HCMV-infected adults. This 'whole ORFeome' approach resulted in the
84 identification of more than 200 new CD4+ and CD8+ T cell epitopes.

85

86

87 **Results**

88 **Targets of HCMV T cell reactivity**

89 To define the epitopes targeted by HCMV-specific T cell responses in healthy
90 adults, we screened PBMCs of 19 subjects, 10 males and 9 females, recruited from the
91 San Diego blood bank (SDBB). The HCMV seropositivity of all the subjects was confirmed
92 by IgG ELISA (**Fig. S1A**). We tested a total of 2593 15-mer HCMV peptides covering a
93 total of 563 ORFs (39). Removing the predicted ORFs that were located entirely within
94 longer ORFs resulted in a set of 359 completely unique ORFs. This set consists of
95 approximately 150 “canonical” ORFs, with an additional 200 identified by ribosomal RNA
96 profiling (39). These 15-mer peptides corresponded to epitopes likely to be dominant
97 based on a bioinformatic method that predicts promiscuous binding to HLA class II
98 molecules (40). Each ORF analyzed contained a minimum of 2 predicted epitopes, with
99 the exception of very small ORFs of less than 15-20 amino acid residues, in which case
100 at least one peptide was synthesized. The 2593 peptides were arranged in 89 pools of
101 28 to 30 15-mers. The PBMC reactivity of each of the 89 pools was assayed directly *ex*
102 *vivo* using an IFN- γ Fluorospot assay. After identifying the pools that resulted in IFN- γ
103 production in HCMV+ individuals, the top 10 most reactive pools (that, on average,
104 accounted for more than 90% of the reactivity observed within each subject) were then
105 deconvoluted to identify the specific epitopes (**Fig. S2**). Representative results from the
106 initial screening and the deconvolution of a pool in a representative subject are shown in
107 **Fig. 1A-B**. In conclusion, the results shown here indicate that human T cell responses to
108 HCMV recognize a wide breadth of different epitope specificities.

109

110 **Characterization of CMV epitope-specific immune responses**

111 The deconvolution of the top 10 pools from each subject identified widespread
112 reactivity directed against 235 unique epitopes (**Fig. S3 and Table 1**). Interestingly,
113 females tended to show both a higher frequency and magnitude of epitope-specific
114 responses when compared to males, although this did not reach statistical significance
115 (**Fig. S4**). On average, each subject recognized 25 epitopes (**Fig. 2A**) and all subjects
116 recognized at least 2 (range 2-57, **Fig. 2B**). Specifically, 6 out of 19 donors recognized
117 21-30 epitopes. A quarter of the epitopes (58 of the 235 recognized) were recognized by
118 three or more subjects (**Fig. 2C**), and these accounted for 76% of the total T cell response
119 (**Fig. 2D**).

120 We further characterized the phenotype of the T cell responses directed against
121 these 58 dominant epitopes by intracellular IFN- γ staining (representative results shown
122 in **Fig. 3A**, with the flow cytometry gating strategy shown in **Fig. S5A**). In the majority of
123 tested subjects, the responding T cells were CD4+. More specifically, 68% of all
124 responding T cells were CD4+ and 13% contained both IFN- γ + CD4+ and CD8+. In 18%
125 of the cases, only CD8+ T cells responded to these 58 epitopes (**Fig. 3B**). Similarly, if the
126 magnitude of the response was considered, 70% of the IFN- γ response was attributable
127 to CD4+ T cells and only 30% emanated from CD8+ T cells (**Fig. 3C**). The fact that the
128 responses were dominated by CD4+ T cells is consistent with the fact that the peptides
129 tested were originally selected based on their predicted likelihood to bind HLA class II
130 alleles. In turn, the occasional identification of epitope-specific CD8+ T cell responses in
131 many cases likely reflects class I epitopes nested within the 15-mer epitopes tested in the

132 screen. Overall, these results indicate that, as expected, the screening strategy employed
133 mostly identifies targets of CD4+ T cell reactivity.

134 **Analysis of the ORF of origin of the identified epitopes**

135 The 235 epitopes identified mapped to a total of 100 of the 359 unique ORFs
136 screened. Of those, 28 ORFs contained >3 immunogenic peptides and 18 ORFs were
137 recognized in 15% or more of the donors (**Fig. 4**). Notably, the previously well-
138 characterized immunodominant ORFs such as envelope glycoprotein B (UL55), IE1
139 (UL123), tegument protein pp65 (UL83), major capsid protein UL86, IE2 (UL122), and
140 pp150 (UL32) were amongst those associated with more than three immunogenic
141 peptides.

142 To address the novelty of our findings, we compared our results with ORFs that
143 have already been reported and curated in the Immune Epitope Database (IEDB
144 <http://www.iedb.org>) (41), as a source of defined epitopes. Specifically, a query of the
145 IEDB in October 2020 for previously characterized targets of T cell responses tested in
146 at least 19 donors and with a minimum response frequency of 15% revealed 7 ORFs that
147 match the conditions of our screening results: UL83/pp65 (ORFL205C), UL123/IE1
148 (ORFL264C), UL122/IE2 (ORFL265C), UL55/gB (ORFL145C), UL32/pp150 (ORFL92C),
149 UL40 (ORFL105C) and UL98 (ORFL229W) (**Fig. 4**).

150 The same query revealed three additional ORFs that were not identified in our
151 screen. These ORFs were associated with a limited number of literature-reported and
152 IEDB curated epitopes: UL75/gH (ORFL184C; 1 epitope), UL44/DNA-pol
153 (ORFL112C.iORF1; 3 epitopes) and UL138 (ORFL313C; 1 epitope). Importantly, our

154 screen identified 93 ORFs that were not previously described as targets of T cell
155 responses (**Fig. 5**).

156 Notably, 52 of these 93 ORFs were already described in the ‘canonical HCMV’
157 annotated genome, but not all have been described as targets of human T cell responses.
158 Even more strikingly, 41 of these 93 ORFs corresponded to those viral mRNAs only
159 identified by recent ribosomal profiling studies (39), providing evidence that they are
160 translated in HCMV infected cells. These results indicate that our approach successfully
161 re-identified known ORFs as targets of T cell responses, and perhaps most importantly,
162 greatly expanded the repertoire of canonical and ‘novel’ ORFs recognized by T cells in
163 healthy adults.

164 **Novel identified epitope pools elicit antigen specific CD4+ T cell responses.**

165 Lastly, we wanted to explore whether the epitopes identified in the presented study
166 could, alone or in combination with previously described epitopes, be utilized to generate
167 epitope “MegaPools” (MP) (42-46) to allow detection of CMV-specific CD4 T cell
168 responses. Accordingly, we generated a ‘P235’ MP encompassing the 235 CMV epitopes
169 identified in the present study. As a comparison, we considered the commercially
170 available CMV peptide pool (Mabtech, catalog 3619-1) encompassing a total of 42 CD4
171 and CD8 epitopes. Additionally, we synthesized a MP of known class II epitopes curated
172 in the IEDB database, encompassing a total of 187 CD4 epitopes (IEDB-II, Table 2).

173 These MPs were tested with PBMC from a new cohort of 20 individuals (6 males
174 and 14 females), which included both HCMV seropositive and seronegative donors (10
175 CMV⁺ and 10 CMV⁻, **Fig. S1B** for IgG ELISA CMV confirmation). None of the PBMC from
176 these subjects were used in the original epitope mapping experiments. PBMCs were

177 stimulated with the Mabtech, P235, IEDB-II, or a combination of both P235/ IEDB-II MPs.
178 CD4+ T cell responses were measured as percentage of activation-induced marker assay
179 positive (OX40+ CD137+) CD4+ T cells and results are displayed in **Fig. 6** (flow cytometry
180 gating strategy shown in **Fig. S5B**).

181 All HCMV MPs tested were associated with significantly higher CD4 AIM
182 responses in HCMV+ individuals compared to HCMV- subjects as shown in **Fig. 5**
183 (statistical differences detailed in figure legend). When comparing AIM responses
184 between the HCMV pools, the P235, IEDB-II and P235/IEDB-II MPs were associated with
185 significantly higher HCMV-specific CD4 responses compared to the Mabtech pool
186 (geometric mean 0.15% vs 0.25% CD4 AIM+, $p=0.01$; and 0.15% vs 0.36%, $p=0.004$,
187 and 0.15% vs 0.46% CD4 AIM+, $p=0.004$, respectively by Wilcoxon test). This was
188 expected, as the Mabtech pool contains fewer epitopes which are also mainly CD8 T cell
189 specific. Additionally, the combination of the P235 and IEDB-II MPs elicited higher CD4
190 responses than either MP alone (geometric mean 0.25% vs 0.46% CD4 AIM+, $p=0.0078$
191 and 0.36% vs 0.46% CD4 AIM+, $p=0.004$, respectively by Wilcoxon test) and had the
192 highest magnitude response of all pools tested. This indicates that the combination of
193 known (IEDB-II MP) and novel epitopes and ORFs (P235 MP) can capture the broadest
194 range of CD4 T-cell responses in HCMV+ individuals, which has high potential for clinical
195 diagnostic use.

196

197

198

199 **Discussion**

200 In this study we have identified >200 new epitopes derived from 100 HCMV ORFs
201 that induce virus-specific T cell responses. Importantly, this demonstrates that the current
202 HLA peptide-binding prediction algorithms that we and others have refined over the last
203 several decades are extremely efficient (47-51), and represent an excellent alternative to
204 synthesizing genome-wide overlapping peptides, especially for large pathogens such as
205 HCMV. Despite the significant diversity in the human HLA repertoire, current advances in
206 algorithm-based epitope identification take into consideration epitopes with potential
207 binding to diverse haplotypes, which undoubtedly contributed to this success (40, 52).
208 Together, this approach allowed us to increase the known T cell epitope landscape for
209 HCMV by greater than 10-fold by synthesizing only 2593 peptides, illustrating both its
210 efficiency and cost effectiveness in deciphering immune targets of large pathogens.

211 We chose to use IFN- γ production as a readout for positive epitope reactivity in a
212 fluorospot-based assay to identify HCMV-specific T cell epitopes in this study. As true for
213 most viral infections, CMV drives a strong Th1-like CD4⁺ response, and most effector
214 and memory viral CD8⁺ T cells also produce this cytokine (53). However, future studies
215 assessing which of these 235 epitopes may elicit HCMV-specific CD4 T cells to produce
216 other cytokines are merited. Previously, we have observed that Dengue virus epitope-
217 specific CD4⁺ T cells can produce both IFN γ and IL-10 (54), something we have also
218 seen during acute CMV infection in mice (55), where IL-10 producing CD4⁺ T cells
219 enhance the duration of viral persistence (56). Recent studies by the Wills and Moss
220 groups show that subsets of HCMV epitope-specific CD4⁺ T cells can produce IL-10 and
221 also display cytolytic markers (57, 58). The potential CTL activity of HCMV-specific CD4⁺

222 T cells has been postulated for many years (59), and our recent results showing that CMV
223 epitope-specific CD4 T cells can directly kill *in vivo* support this hypothesis (60). Taken
224 together, our identification of >200 new T cell epitopes that elicit IFN γ production in this
225 study provide us and others in the field valuable new tools to dissect the phenotypes and
226 effector functions of HCMV-specific CD4 T cells in cases of both healthy and immune
227 compromised patients, and will also help instruct ongoing vaccine efforts.

228 Of the 100 ORFs which we show here to be sources of specific T cell epitopes, 41
229 were uniquely identified as ribosome-bound RNAs in HCMV infected fibroblasts (39), with
230 these 41 yielding 50 unique epitopes. Notably, of these 41 ORFs, 17 are predicted to
231 produce proteins <50 amino acids in length, and 7 contain non-ATG start codons. This is
232 consistent with recent studies suggesting that the short/‘cryptic’ mRNAs present in both
233 virally infected and tumor cells can be translated, proteolytically processed and loaded
234 onto HLA molecules, resulting in the induction of epitope-specific T cell responses (61-
235 63). Interestingly, one of the larger 41 ORFs that contains two newly identified T cell
236 epitopes (ORFL147C, 476 amino acids) has very recently been shown to regulate RNA
237 binding/processing, and its deletion compromises CMV replication in fibroblasts (64).
238 Despite >20% of the novel T cell epitopes identified here being derived from these newly
239 described, ribosome-associated HCMV RNAs, no more than 2 of the 19 healthy donors
240 analyzed produce T cells specific for any single one of these epitopes. This indicates that
241 these novel ORFs 1) may not be broad targets of T cell responses in infected persons, 2)
242 that specific individuals may more efficiently present epitopes derived from short/cryptic
243 HCMV RNAs or 3) that minor HLA molecules may present them, with other possibilities
244 also existing. Additionally, whether the proteins derived from these short ORFs are stable

245 and play a role in the HCMV lifecycle remains an open question. Finally, we also identified
246 24 epitopes derived from 14 'canonical' HCMV ORFs where the only historic support for
247 their existence was the presence of their RNA in infected cells or bioinformatic analyses.
248 Notably, a recent comprehensive study where 169 predicted canonical HCMV proteins
249 (including these 14) were epitope-tagged, expressed stably in infected cells,
250 immunoprecipitated and analyzed for interacting proteins by mass spectrometry supports
251 our results that these ORFs are expressed as proteins (64).

252 Of the 59 canonical ORFs that we have identified here to contain T cell epitopes,
253 >25% of these are known to function as immunomodulatory proteins (65). This is
254 intriguing, as perhaps these HCMV proteins are more subject to being localized to
255 antigen-processing or presentation compartments within infected cells. One of these
256 epitopes is derived from the HCMV IL-10 orthologue, which is being considered as a
257 potential HCMV vaccine candidate (66, 67). Additionally, 3 epitopes were found to be
258 embedded within the viral UL128 protein, a critical component of the pentameric envelope
259 protein complex (UL128-131/gH/gL) that mediates entry of HCMV into non-fibroblast cell
260 types (68, 69). This is also of high potential interest in the context of vaccine development,
261 as many believe the pentamer should be included in a viral- or subunit-based approach
262 (70). Notably, both vIL-10 and UL128 have largely been considered only in the context of
263 their abilities to induce antibody-based vaccine protection, but our identification of T cell
264 epitopes derived from both these HCMV proteins suggests they may function to prime
265 both humoral and cellular immunity.

266

267 **Methods**

268 **Study design**

269 For the initial CMV ORF screen, the responses of 19 CMV-seropositive subjects
270 were evaluated. PBMCs were stimulated with 89 pools covering 563 ORFs of HCMV.
271 Each pool comprised of 28-30 15-mer peptides overlapping by 10 residues. PBMCs that
272 were found reactive to a pool were further tested against individual peptides contained in
273 the pool using IFN- γ Fluorospot assay. Flow cytometry was then used to further
274 characterize the epitopes recognized by PBMCs stimulated with individual peptides by
275 detecting IFN- γ production from CD8+ and CD4+ T cells.

276 For the CMV-235 validation and comparison screen, the responses of a new cohort
277 consisting of 10 CMV-seropositive and 10 seronegative subjects were evaluated. PBMCs
278 were stimulated with CMV-Mabtech peptide pool (Catalog 3619-1), CMV-IEDB peptide
279 pool (Table 2) (44, 46), CMV-235 pool, or a combination of both CMV-IEDB and CMV-
280 235 pools. PBMC responses were assayed using the same IFN- γ Fluorospot assay.
281 These studies were approved by the institutional review board committee at La Jolla
282 Institute protocol number: VD-112 and VD-174.

283 **Subjects**

284 19 subjects (10 males and 9 females) were recruited anonymously from San Diego
285 blood bank (SDBB) for the initial CMV ORF screens. For the CMV-235 comparison
286 screens, samples from 20 subjects (6 males and 14 females) were obtained by La Jolla
287 Institute Clinical Core and Continental Services Group (Miami, FL) for prior, unrelated
288 studies. Blood samples were collected by trained staff. At the time of enrollment in the
289 initial studies, all individual subjects provided informed consent that any leftover sample
290 could be used for future studies, which includes this study. These subjects were

291 considered healthy as defined by no known history of any significant systemic diseases
292 (not limited to autoimmune disease, diabetes, kidney or liver disease, congestive heart
293 failure, malignancy, coagulopathy, hepatitis B or C, or HIV). The demographics of those
294 subjects are provided in **Table 3**.

295 The IgG antibodies of the subjects for both cohorts were measured using
296 Cytomegalovirus IgG Elisa kit from Genway Biotech Inc. according to manufacturer's
297 instructions.

298 **Peptide prediction**

299 Based on the 7-allele method as previously described (40), 2593 peptides were
300 predicted for 563 potential HCMV ORFs. Of the 751 ORFs predicted by ribosomal
301 profiling (39), those smaller than 15 amino acids were excluded, and only one peptide of
302 ORFs 15-20 amino acids in length were selected for screening.

303 **Peptide libraries and pool preparation**

304 The predicted peptides were commercially synthesized as crude material by TC
305 Peptide Lab (www.tcpeptidelab.com; San Diego, CA). The peptides were solubilized in
306 DMSO at a concentration of 20 mg/ml and spot checked for quality by mass spectrometry.
307 The peptides were pooled into peptide pools containing 28-30 peptides constituting
308 multiple ORFs per pool. A total of 89 pools were prepared covering 563 ORFs of HCMV.
309 The final concentration of each pool was 0.7 mg/ml.

310 For the IEDB-II (Table 2) and P235 (Table 1) peptide pools peptides were
311 synthesized by A&A Ltd, San Diego, resuspended in DMSO, pooled and sequentially
312 lyophilized as previously described (71). The IEDB-II peptide pool was developed based
313 on data available in the IEDB (www.iedb.org) (41). The MHC class II restricted epitopes

314 for CMV was extracted from the IEDB in October of 2020 using the following query;
315 Organism: human herpesvirus 5 (ID:10359), positive assays only, no B cell assays, MHC
316 restriction type: class II, host: *Homo sapiens*. The resulting 187 epitopes (table 2) were
317 filtered for size (13-20 amino acids) and discovered using one of the following assays:
318 ELISPOT, ICS, multi- or tetramers, proliferation and “helper response”. The CMV peptide
319 pool for human CD4 and CD8 T cells containing 42 peptides (14 MHC class II restricted
320 and 28 MHC class I restricted) representing pp50, pp65, IE1, IE2, and envelope
321 glycoprotein B was purchased from Mabtech.

322 **Isolation of PBMC by Ficoll-Paque density gradient centrifugation**

323 One-unit blood from each donor was processed for PBMC isolation. Briefly, blood
324 was centrifuged and the top layer of plasma was removed. The remaining blood was
325 diluted and layered over 15 ml of Ficoll-Paque. Tubes were spun at room temperature in
326 a swinging bucket rotor without brake applied. The PBMC interface was carefully removed
327 by pipetting and washed with PBS by centrifugation at 800 rpm for 10 mins with brakes
328 off. PBMC pellet was resuspended in RPMI media, cell number and viability were
329 determined by trypan blue staining and cells were cryopreserved in liquid nitrogen in
330 freezing media (90% Fetal bovine serum and 10% DMSO) at a density of 30 million/ml
331 and stored until further processed.

332 **Fluorospot assay**

333 PBMC were thawed, washed and counted for viability using the trypan blue
334 exclusion method. 200,000 cells were plated in triplicates and stimulated with pools
335 (2µg/ml) or peptides (10µg/ml), PHA (10µg/ml) or medium containing equivalent amount
336 of DMSO in 96- well plates (Immubilion-P, Millipore) previously coated with anti IFN-γ

337 antibody (1-D1K, Mabtech, Stockholm, Sweden). After 20 hr incubation at 37°C, cells
338 were discarded and wells were washed six times with PBS/0.05% Tween 20 using an
339 automated plate washer and further incubated with IFN- γ antibody (7-B6-1-FS-BAM) for
340 2 hrs at room temperature. After incubation, wells were washed and incubated with
341 fluorophore conjugated anti-BAM-490 antibody for 1 hr at room temperature. Finally, the
342 plates were washed and incubated with fluorescence enhancer for 15 min, blotted dry
343 and fluorescent spots were counted by computer assisted image analysis (IRIS
344 Fluorospot reader, Mabtech, Sweden).

345 Each pool or peptide was considered positive compared to the background that
346 had equivalent amount of DMSO based on the following criteria: (i) 20 or more spot
347 forming cells (SFC) per 10^6 PBMC after background subtraction, (ii) the stimulation index
348 greater than 2, and (iii) $p < 0.05$ by student's t test or Poisson distribution test when
349 comparing the peptide or pool triplicates with the negative control triplicate.

350 **Intracellular cytokine assay for IFN- γ**

351 Intracellular staining for IFN- γ and flow cytometry was performed to detect antigen
352 specific T cell responses. 1×10^6 PBMCs suspended in RPMI medium supplemented with
353 1-% heat inactivated human AB serum, glutamine and penicillin streptomycin were plated
354 in U-bottom 96 well plates. After overnight resting at 37°C, PBMCs were spun and
355 replaced with fresh RPMI media and stimulated with individual peptides at a concentration
356 of 10 $\mu\text{g/ml}$. PHA at a concentration of 5 $\mu\text{g/ml}$ was used as a positive control. After 1 hr
357 of incubation at 37°C, 2 $\mu\text{g/ml}$ of Brefeldin was added and cell were further incubated at
358 37°C for additional 5 hrs. The cells were then harvested, washed with 200 μl of MACS
359 Buffer and stained with a cocktail of antibodies that contained CD3-Af700 (eBioscience,

360 clone UCHT1), CD4-APCef780 (eBioscience, clone RPA-T4), CD8-BV650 (Biolegend,
361 clone RPA-T8), CD14-V500 (BD Biosciences, clone M5E2), CD19-V500 (BD
362 Biosciences, clone HIB19), and fixable viability dye-e506 for 30 min at 4°C. The cells
363 were then washed thrice with 200 µl MACS buffer, fixed using 4% PFA for 10 mins at 4°C,
364 washed with 200 µl PBS and rested at 4°C overnight in 200 µl MACS buffer. The following
365 day, cells were washed, permeabilized by washing with 200 µl saponin buffer (0.5 %
366 saponin in PBS), washed with blocking buffer (10% human serum prepared in saponin
367 buffer) and stained with IFN- γ -FITC (eBioscience, clone 4S.B3) antibody at room
368 temperature for 30 mins. The cells were finally washed with PBS and suspended in 200
369 µl PBS.

370 The cells were acquired on ZE5 Biorad plate reader and further analysis was done
371 on FlowJo software. Gates were applied on live single cells for CD3+, CD4+ and CD8+ T
372 cell populations. The percentage of reactive CD4+ or CD8+ IFN- γ T cells were expressed
373 as a percent of the total number of parent population analyzed. Reactive populations met
374 the following 2 criteria: (i) well-defined cell population positive for both IFN- γ and CD4 or
375 CD8 constituting at least 0.02% (post subtracting their corresponding DMSO controls) of
376 the total number of CD4+ or CD8+ cells analyzed (ii) stimulation index greater than 2.

377 **Activation induced marker (AIM) assay**

378 PBMC were thawed, washed and counted for viability using the trypan blue
379 exclusion method. 1 million cells per donor/condition were plated and cultured in the
380 presence of the CMV specific pools (1µg/mL for P235 and IEDB-II, 2µg/mL for Mabtech
381 pool), PHA (10µg/mL), or medium containing equivalent amount of DMSO in 96-well U-
382 bottom plates. Cells were then harvested, washed with 200µl of MACS Buffer and stained

383 with a cocktail of antibodies that contained CD3-Af700 (eBioscience, clone UCHT1), CD4-
384 BV605 (eBioscience , clone RPA-T4), CD8-PerCP-Cy5.5 (Biolegend, clone HIT8a),
385 CD14-V500 (BD Biosciences, clone M5E2), CD19-V500 (BD Biosciences, clone HIB19),
386 OX40-PE-Cy7 (Ber-ACT35), CD137-APC (4B4-1), and fixable viability dye-e506 for 30
387 min at 4°C. The cells were then washed thrice with 200 µl MACS buffer, fixed using 4%
388 PFA for 10 mins at 4°C, and resuspended in 200 µl of PBS for acquisition.

389 Cells were acquired on a BD LSRFortessa and further analysis was done on
390 FlowJo software. As previously described (44, 72), quantification of live, singlet antigen
391 specific CD4 T cells was determined as a percentage of their OX40+CD137+ expression
392 (AIM+). CMV specific AIM+ CD4 T cell signals were background subtracted with their
393 corresponding negative control DMSO samples, with a minimal DMSO level set to
394 0.005%. The limit of detection (LOD) for the AIM+ assay was calculated by multiplying
395 the upper confidence interval of the geometric mean of all DMSO samples by 2 (0.03).

396 **Statistical analysis**

397 Statistical analyses were performed using GraphPad Prism versions 8.1.1 and
398 8.4.3. Statistical details are provided with each figure.

399

400

401

402

403

404

405

406 **Acknowledgement**

407 We would also like to thank all donors that participated in the study. We also thank the
408 La Jolla Institute for Immunology Clinical Studies Group and Flow Cytometry Core for all
409 the invaluable help.

410

411 **Author contribution**

412 AS and CAB conceived the study. RD and GW performed the experiments and
413 analyzed the data. SKD performed peptide prediction, JP processed blood samples,
414 RD, GW, and GP conducted ELISA, JS helped with the quality check of synthesized
415 peptides, AG designed the IEDB-II pool, CLA, AS, CAB directed the study, RD, AS,
416 CAB wrote the manuscript taking input from other authors.

417

418 **Funding**

419 This work was supported by NIH Grants AI139749 and AI101423 to C.A.B, and NIH
420 contracts 75N93019C00065 to A.S. and 75N93019C00067 to C.L.A.

421

422 **Conflict of interest**

423 The authors declare that they have no conflict of interest.

424

425

426

427

428

429

430 **References**

- 431 1. Collins-McMillen D, Buehler J, Peppenelli M, Goodrum F. 2018. Molecular
432 Determinants and the Regulation of Human Cytomegalovirus Latency and Reactivation.
433 *Viruses* 10.
- 434 2. Hargett D, Shenk TE. 2010. Experimental human cytomegalovirus latency in CD14+
435 monocytes. *Proc Natl Acad Sci U S A* 107:20039-44.
- 436 3. Chaturvedi S, Klein J, Vardi N, Bolovan-Fritts C, Wolf M, Du K, Mlera L, Calvert M,
437 Moorman NJ, Goodrum F, Huang B, Weinberger LS. 2020. A molecular mechanism for
438 probabilistic bet hedging and its role in viral latency. *Proc Natl Acad Sci U S A*
439 117:17240-17248.
- 440 4. Davison AJ, Dolan A, Akter P, Addison C, Dargan DJ, Alcendor DJ, McGeoch DJ,
441 Hayward GS. 2003. The human cytomegalovirus genome revisited: comparison with the
442 chimpanzee cytomegalovirus genome. *J Gen Virol* 84:17-28.
- 443 5. Mattes FM, Vargas A, Kopycinski J, Hainsworth EG, Sweny P, Nebbia G, Bazeos A,
444 Lowdell M, Klenerman P, Phillips RE, Griffiths PD, Emery VC. 2008. Functional
445 impairment of cytomegalovirus specific CD8 T cells predicts high-level replication after
446 renal transplantation. *Am J Transplant* 8:990-9.
- 447 6. Walker S, Fazou C, Crough T, Holdsworth R, Kiely P, Veale M, Bell S, Gailbraith A,
448 McNeil K, Jones S, Khanna R. 2007. Ex vivo monitoring of human cytomegalovirus-
449 specific CD8+ T-cell responses using QuantiFERON-CMV. *Transpl Infect Dis* 9:165-70.
- 450 7. Clarke LM, Duerr A, Feldman J, Sierra MF, Daidone BJ, Landesman SH. 1996. Factors
451 associated with cytomegalovirus infection among human immunodeficiency virus type 1-
452 seronegative and -seropositive women from an urban minority community. *J Infect Dis*
453 173:77-82.
- 454 8. Doyle M, Atkins JT, Rivera-Matos IR. 1996. Congenital cytomegalovirus infection in
455 infants infected with human immunodeficiency virus type 1. *Pediatr Infect Dis J* 15:1102-
456 6.
- 457 9. Duryea EL, Sanchez PJ, Sheffield JS, Jackson GL, Wendel GD, McElwee BS, Boney LF,
458 Mallory MM, Owen KE, Stehel EK. 2010. Maternal human immunodeficiency virus
459 infection and congenital transmission of cytomegalovirus. *Pediatr Infect Dis J* 29:915-8.
- 460 10. Kovacs A, Schluchter M, Easley K, Demmler G, Shearer W, La Russa P, Pitt J, Cooper
461 E, Goldfarb J, Hodes D, Kattan M, McIntosh K. 1999. Cytomegalovirus infection and
462 HIV-1 disease progression in infants born to HIV-1-infected women. *Pediatric*
463 *Pulmonary and Cardiovascular Complications of Vertically Transmitted HIV Infection*
464 *Study Group. N Engl J Med* 341:77-84.
- 465 11. Schoenfisch AL, Dollard SC, Amin M, Gardner LI, Klein RS, Mayer K, Rompalo A,
466 Sobel JD, Cannon MJ. 2011. Cytomegalovirus (CMV) shedding is highly correlated with
467 markers of immunosuppression in CMV-seropositive women. *J Med Microbiol* 60:768-
468 774.
- 469 12. Demmler-Harrison GJ. 2009. Congenital cytomegalovirus: Public health action towards
470 awareness, prevention, and treatment. *J Clin Virol* 46 Suppl 4:S1-5.

- 471 13. Jeon J, Victor M, Adler SP, Arwady A, Demmler G, Fowler K, Goldfarb J, Keyserling H,
472 Massoudi M, Richards K, Staras SA, Cannon MJ. 2006. Knowledge and awareness of
473 congenital cytomegalovirus among women. *Infect Dis Obstet Gynecol* 2006:80383.
- 474 14. Fowler KB, Stagno S, Pass RF, Britt WJ, Boll TJ, Alford CA. 1992. The outcome of
475 congenital cytomegalovirus infection in relation to maternal antibody status. *N Engl J*
476 *Med* 326:663-7.
- 477 15. Ross SA, Fowler KB, Ashrith G, Stagno S, Britt WJ, Pass RF, Boppana SB. 2006.
478 Hearing loss in children with congenital cytomegalovirus infection born to mothers with
479 preexisting immunity. *J Pediatr* 148:332-6.
- 480 16. Fowler KB, Boppana SB. 2006. Congenital cytomegalovirus (CMV) infection and
481 hearing deficit. *J Clin Virol* 35:226-31.
- 482 17. Ross SA, Boppana SB. 2005. Congenital cytomegalovirus infection: outcome and
483 diagnosis. *Semin Pediatr Infect Dis* 16:44-9.
- 484 18. Britt WJ. 2018. Maternal Immunity and the Natural History of Congenital Human
485 Cytomegalovirus Infection. *Viruses* 10.
- 486 19. Griffiths PD. 2002. Strategies to prevent CMV infection in the neonate. *Semin Neonatol*
487 7:293-9.
- 488 20. Kimberlin DW, Lin CY, Sanchez PJ, Demmler GJ, Dankner W, Shelton M, Jacobs RF,
489 Vaudry W, Pass RF, Kiell JM, Soong SJ, Whitley RJ, National Institute of A, Infectious
490 Diseases Collaborative Antiviral Study G. 2003. Effect of ganciclovir therapy on hearing
491 in symptomatic congenital cytomegalovirus disease involving the central nervous system:
492 a randomized, controlled trial. *J Pediatr* 143:16-25.
- 493 21. Michaels MG, Greenberg DP, Sabo DL, Wald ER. 2003. Treatment of children with
494 congenital cytomegalovirus infection with ganciclovir. *Pediatr Infect Dis J* 22:504-9.
- 495 22. Whitley RJ, Cloud G, Gruber W, Storch GA, Demmler GJ, Jacobs RF, Dankner W,
496 Spector SA, Starr S, Pass RF, Stagno S, Britt WJ, Alford C, Jr., Soong S, Zhou XJ,
497 Sherrill L, FitzGerald JM, Sommadossi JP. 1997. Ganciclovir treatment of symptomatic
498 congenital cytomegalovirus infection: results of a phase II study. National Institute of
499 Allergy and Infectious Diseases Collaborative Antiviral Study Group. *J Infect Dis*
500 175:1080-6.
- 501 23. Benedict CA. 2013. A CMV vaccine: TREATing despite the TRICKs. *Expert Rev*
502 *Vaccines* 12:1235-7.
- 503 24. Permar SR, Schleiss MR, Plotkin SA. 2018. Advancing Our Understanding of Protective
504 Maternal Immunity as a Guide for Development of Vaccines To Reduce Congenital
505 Cytomegalovirus Infections. *J Virol* 92.
- 506 25. Plotkin SA, Boppana SB. 2019. Vaccination against the human cytomegalovirus. *Vaccine*
507 37:7437-7442.
- 508 26. Li CR, Greenberg PD, Gilbert MJ, Goodrich JM, Riddell SR. 1994. Recovery of HLA-
509 restricted cytomegalovirus (CMV)-specific T-cell responses after allogeneic bone
510 marrow transplant: correlation with CMV disease and effect of ganciclovir prophylaxis.
511 *Blood* 83:1971-9.
- 512 27. Quinnan GV, Jr., Kirmani N, Rook AH, Manischewitz JF, Jackson L, Moreschi G,
513 Santos GW, Saral R, Burns WH. 1982. Cytotoxic t cells in cytomegalovirus infection:
514 HLA-restricted T-lymphocyte and non-T-lymphocyte cytotoxic responses correlate with
515 recovery from cytomegalovirus infection in bone-marrow-transplant recipients. *N Engl J*
516 *Med* 307:7-13.

- 517 28. Reusser P, Riddell SR, Meyers JD, Greenberg PD. 1991. Cytotoxic T-lymphocyte
518 response to cytomegalovirus after human allogeneic bone marrow transplantation: pattern
519 of recovery and correlation with cytomegalovirus infection and disease. *Blood* 78:1373-
520 80.
- 521 29. Smith CJ, Quinn M, Snyder CM. 2016. CMV-Specific CD8 T Cell Differentiation and
522 Localization: Implications for Adoptive Therapies. *Front Immunol* 7:352.
- 523 30. Sylwester AW, Mitchell BL, Edgar JB, Taormina C, Pelte C, Ruchti F, Sleath PR,
524 Grabstein KH, Hosken NA, Kern F, Nelson JA, Picker LJ. 2005. Broadly targeted human
525 cytomegalovirus-specific CD4+ and CD8+ T cells dominate the memory compartments
526 of exposed subjects. *J Exp Med* 202:673-85.
- 527 31. Kern F, Bunde T, Faulhaber N, Kiecker F, Khatamzas E, Rudawski IM, Pruss A,
528 Gratama JW, Volkmer-Engert R, Ewert R, Reinke P, Volk HD, Picker LJ. 2002.
529 Cytomegalovirus (CMV) phosphoprotein 65 makes a large contribution to shaping the T
530 cell repertoire in CMV-exposed individuals. *J Infect Dis* 185:1709-16.
- 531 32. Kern F, Surel IP, Faulhaber N, Frommel C, Schneider-Mergener J, Schonemann C,
532 Reinke P, Volk HD. 1999. Target structures of the CD8(+)-T-cell response to human
533 cytomegalovirus: the 72-kilodalton major immediate-early protein revisited. *J Virol*
534 73:8179-84.
- 535 33. Khan N, Best D, Bruton R, Nayak L, Rickinson AB, Moss PA. 2007. T cell recognition
536 patterns of immunodominant cytomegalovirus antigens in primary and persistent
537 infection. *J Immunol* 178:4455-65.
- 538 34. Khan N, Cobbold M, Keenan R, Moss PA. 2002. Comparative analysis of CD8+ T cell
539 responses against human cytomegalovirus proteins pp65 and immediate early 1 shows
540 similarities in precursor frequency, oligoclonality, and phenotype. *J Infect Dis* 185:1025-
541 34.
- 542 35. Elkington R, Walker S, Crough T, Menzies M, Tellam J, Bharadwaj M, Khanna R. 2003.
543 Ex vivo profiling of CD8+-T-cell responses to human cytomegalovirus reveals broad and
544 multispecific reactivities in healthy virus carriers. *J Virol* 77:5226-40.
- 545 36. Elkington R, Shoukry NH, Walker S, Crough T, Fazou C, Kaur A, Walker CM, Khanna
546 R. 2004. Cross-reactive recognition of human and primate cytomegalovirus sequences by
547 human CD4 cytotoxic T lymphocytes specific for glycoprotein B and H. *Eur J Immunol*
548 34:3216-26.
- 549 37. Paston SJ, Dodi IA, Madrigal JA. 2004. Progress made towards the development of a
550 CMV peptide vaccine. *Hum Immunol* 65:544-9.
- 551 38. Fuhrmann S, Streitz M, Reinke P, Volk HD, Kern F. 2008. T cell response to the
552 cytomegalovirus major capsid protein (UL86) is dominated by helper cells with a large
553 polyfunctional component and diverse epitope recognition. *J Infect Dis* 197:1455-8.
- 554 39. Stern-Ginossar N, Weisburd B, Michalski A, Le VT, Hein MY, Huang SX, Ma M, Shen
555 B, Qian SB, Hengel H, Mann M, Ingolia NT, Weissman JS. 2012. Decoding human
556 cytomegalovirus. *Science* 338:1088-93.
- 557 40. Paul S, Lindestam Arlehamn CS, Scriba TJ, Dillon MB, Oseroff C, Hinz D, McKinney
558 DM, Carrasco Pro S, Sidney J, Peters B, Sette A. 2015. Development and validation of a
559 broad scheme for prediction of HLA class II restricted T cell epitopes. *J Immunol*
560 Methods 422:28-34.

- 561 41. Vita R, Mahajan S, Overton JA, Dhanda SK, Martini S, Cantrell JR, Wheeler DK, Sette
562 A, Peters B. 2019. The Immune Epitope Database (IEDB): 2018 update. *Nucleic Acids*
563 *Res* 47:D339-D343.
- 564 42. Bancroft T, Dillon MB, da Silva Antunes R, Paul S, Peters B, Crotty S, Lindestam
565 Arlehamn CS, Sette A. 2016. Th1 versus Th2 T cell polarization by whole-cell and
566 acellular childhood pertussis vaccines persists upon re-immunization in adolescence and
567 adulthood. *Cell Immunol* 304-305:35-43.
- 568 43. da Silva Antunes R, Paul S, Sidney J, Weiskopf D, Dan JM, Phillips E, Mallal S, Crotty
569 S, Sette A, Lindestam Arlehamn CS. 2017. Definition of Human Epitopes Recognized in
570 Tetanus Toxoid and Development of an Assay Strategy to Detect Ex Vivo Tetanus CD4+
571 T Cell Responses. *PLoS One* 12:e0169086.
- 572 44. Grifoni A, Weiskopf D, Ramirez SI, Mateus J, Dan JM, Moderbacher CR, Rawlings SA,
573 Sutherland A, Premkumar L, Jadi RS, Marrama D, de Silva AM, Frazier A, Carlin AF,
574 Greenbaum JA, Peters B, Krammer F, Smith DM, Crotty S, Sette A. 2020. Targets of T
575 Cell Responses to SARS-CoV-2 Coronavirus in Humans with COVID-19 Disease and
576 Unexposed Individuals. *Cell* 181:1489-1501 e15.
- 577 45. Lindestam Arlehamn CS, McKinney DM, Carpenter C, Paul S, Rozot V, Makgotlho E,
578 Gregg Y, van Rooyen M, Ernst JD, Hatherill M, Hanekom WA, Peters B, Scriba TJ,
579 Sette A. 2016. A Quantitative Analysis of Complexity of Human Pathogen-Specific CD4
580 T Cell Responses in Healthy M. tuberculosis Infected South Africans. *PLoS Pathog*
581 12:e1005760.
- 582 46. Mateus J, Grifoni A, Tarke A, Sidney J, Ramirez SI, Dan JM, Burger ZC, Rawlings SA,
583 Smith DM, Phillips E, Mallal S, Lammers M, Rubiro P, Quiambao L, Sutherland A, Yu
584 ED, da Silva Antunes R, Greenbaum J, Frazier A, Markmann AJ, Premkumar L, de Silva
585 A, Peters B, Crotty S, Sette A, Weiskopf D. 2020. Selective and cross-reactive SARS-
586 CoV-2 T cell epitopes in unexposed humans. *Science* 370:89-94.
- 587 47. Nielsen M, Justesen S, Lund O, Lundegaard C, Buus S. 2010. NetMHCIIpan-2.0 -
588 Improved pan-specific HLA-DR predictions using a novel concurrent alignment and
589 weight optimization training procedure. *Immunome Res* 6:9.
- 590 48. Nielsen M, Lund O. 2009. NN-align. An artificial neural network-based alignment
591 algorithm for MHC class II peptide binding prediction. *BMC Bioinformatics* 10:296.
- 592 49. Paul S, Kolla RV, Sidney J, Weiskopf D, Fleri W, Kim Y, Peters B, Sette A. 2013.
593 Evaluating the immunogenicity of protein drugs by applying in vitro MHC binding data
594 and the immune epitope database and analysis resource. *Clin Dev Immunol* 2013:467852.
- 595 50. Salimi N, Fleri W, Peters B, Sette A. 2012. The immune epitope database: a historical
596 retrospective of the first decade. *Immunology* 137:117-23.
- 597 51. Wang P, Sidney J, Kim Y, Sette A, Lund O, Nielsen M, Peters B. 2010. Peptide binding
598 predictions for HLA DR, DP and DQ molecules. *BMC Bioinformatics* 11:568.
- 599 52. Dhanda SK, Karosiene E, Edwards L, Grifoni A, Paul S, Andreatta M, Weiskopf D,
600 Sidney J, Nielsen M, Peters B, Sette A. 2018. Predicting HLA CD4 Immunogenicity in
601 Human Populations. *Front Immunol* 9:1369.
- 602 53. Lim EY, Jackson SE, Wills MR. 2020. The CD4+ T Cell Response to Human
603 Cytomegalovirus in Healthy and Immunocompromised People. *Front Cell Infect*
604 *Microbiol* 10:202.
- 605 54. Tian Y, Seumois G, De-Oliveira-Pinto LM, Mateus J, Herrera-de la Mata S, Kim C, Hinz
606 D, Goonawardhana NDS, de Silva AD, Premawansa S, Premawansa G, Wijewickrama A,

- 607 Balmaseda A, Grifoni A, Vijayanand P, Harris E, Peters B, Sette A, Weiskopf D. 2019.
608 Molecular Signatures of Dengue Virus-Specific IL-10/IFN-gamma Co-producing CD4 T
609 Cells and Their Association with Dengue Disease. *Cell Rep* 29:4482-4495 e4.
- 610 55. Wehrens EJ, Wong KA, Gupta A, Khan A, Benedict CA, Zuniga EI. 2018. IL-27
611 regulates the number, function and cytotoxic program of antiviral CD4 T cells and
612 promotes cytomegalovirus persistence. *PLoS One* 13:e0201249.
- 613 56. Humphreys IR, de Trez C, Kinkade A, Benedict CA, Croft M, Ware CF. 2007.
614 Cytomegalovirus exploits IL-10-mediated immune regulation in the salivary glands. *J*
615 *Exp Med* 204:1217-25.
- 616 57. Jackson SE, Sedikides GX, Mason GM, Okecha G, Wills MR. 2017. Human
617 Cytomegalovirus (HCMV)-Specific CD4(+) T Cells Are Polyfunctional and Can
618 Respond to HCMV-Infected Dendritic Cells In Vitro. *J Virol* 91.
- 619 58. Pachnio A, Ciauriz M, Begum J, Lal N, Zuo J, Beggs A, Moss P. 2016.
620 Cytomegalovirus Infection Leads to Development of High Frequencies of Cytotoxic
621 Virus-Specific CD4+ T Cells Targeted to Vascular Endothelium. *PLoS Pathog*
622 12:e1005832.
- 623 59. van Leeuwen EM, Remmerswaal EB, Vossen MT, Rowshani AT, Wertheim-van Dillen
624 PM, van Lier RA, ten Berge IJ. 2004. Emergence of a CD4+CD28- granzyme B+,
625 cytomegalovirus-specific T cell subset after recovery of primary cytomegalovirus
626 infection. *J Immunol* 173:1834-41.
- 627 60. Verma S, Weiskopf D, Gupta A, McDonald B, Peters B, Sette A, Benedict CA. 2016.
628 Cytomegalovirus-Specific CD4 T Cells Are Cytolytic and Mediate Vaccine Protection. *J*
629 *Virol* 90:650-8.
- 630 61. Erhard F, Dolken L, Schilling B, Schlosser A. 2020. Identification of the Cryptic HLA-I
631 Immunopeptidome. *Cancer Immunol Res* 8:1018-1026.
- 632 62. Erhard F, Halenius A, Zimmermann C, L'Hernault A, Kowalewski DJ, Weekes MP,
633 Stevanovic S, Zimmer R, Dolken L. 2018. Improved Ribo-seq enables identification of
634 cryptic translation events. *Nat Methods* 15:363-366.
- 635 63. Laumont CM, Perreault C. 2018. Exploiting non-canonical translation to identify new
636 targets for T cell-based cancer immunotherapy. *Cell Mol Life Sci* 75:607-621.
- 637 64. Nobre LV, Nightingale K, Ravenhill BJ, Antrobus R, Soday L, Nichols J, Davies JA,
638 Seirafian S, Wang EC, Davison AJ, Wilkinson GW, Stanton RJ, Huttlin EL, Weekes MP.
639 2019. Human cytomegalovirus interactome analysis identifies degradation hubs, domain
640 associations and viral protein functions. *Elife* 8.
- 641 65. Picarda G, Benedict CA. 2018. Cytomegalovirus: Shape-Shifting the Immune System. *J*
642 *Immunol* 200:3881-3889.
- 643 66. Deere JD, Chang WLW, Villalobos A, Schmidt KA, Deshpande A, Castillo LD, Fike J,
644 Walter MR, Barry PA, Hartigan-O'Connor DJ. 2019. Neutralization of rhesus
645 cytomegalovirus IL-10 reduces horizontal transmission and alters long-term immunity.
646 *Proc Natl Acad Sci U S A* 116:13036-13041.
- 647 67. Eberhardt MK, Deshpande A, Chang WL, Barthold SW, Walter MR, Barry PA. 2013.
648 Vaccination against a virus-encoded cytokine significantly restricts viral challenge. *J*
649 *Virol* 87:11323-31.
- 650 68. Nguyen CC, Kamil JP. 2018. Pathogen at the Gates: Human Cytomegalovirus Entry and
651 Cell Tropism. *Viruses* 10.

- 652 69. Vanarsdall AL, Johnson DC. 2012. Human cytomegalovirus entry into cells. *Curr Opin*
653 *Virology* 2:37-42.
- 654 70. Gerna G, Lilleri D. 2019. Human cytomegalovirus (HCMV) infection/re-infection:
655 development of a protective HCMV vaccine. *New Microbiol* 42:1-20.
- 656 71. Carrasco Pro S, Sidney J, Paul S, Lindestam Arlehamn C, Weiskopf D, Peters B, Sette A.
657 2015. Automatic Generation of Validated Specific Epitope Sets. *J Immunol Res*
658 2015:763461.
- 659 72. Dan JM, Mateus J, Kato Y, Hastie KM, Yu ED, Faliti CE, Grifoni A, Ramirez SI, Haupt
660 S, Frazier A, Nakao C, Rayaprolu V, Rawlings SA, Peters B, Krammer F, Simon V,
661 Saphire EO, Smith DM, Weiskopf D, Sette A, Crotty S. 2021. Immunological memory to
662 SARS-CoV-2 assessed for up to 8 months after infection. *Science* 371.
663

664

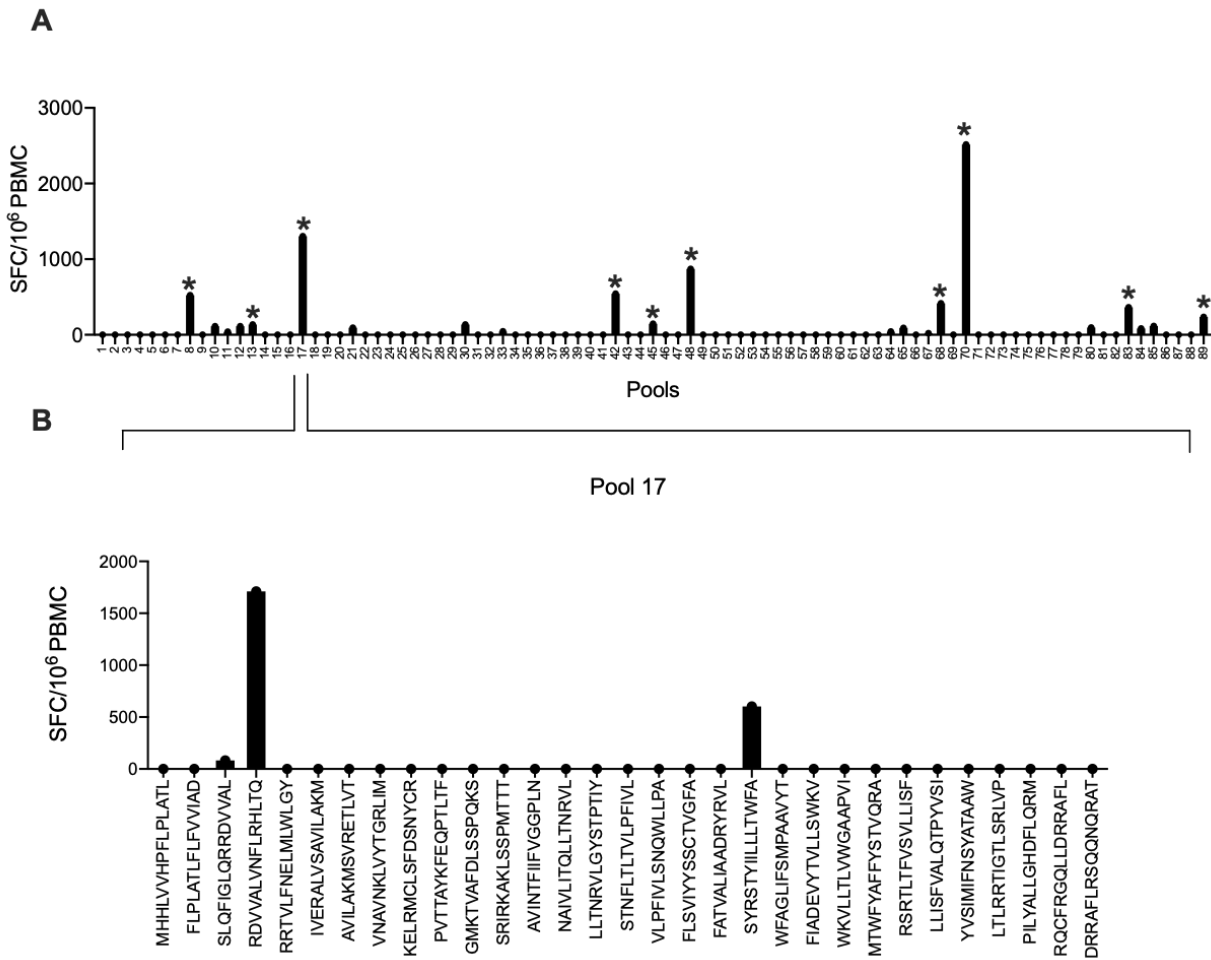
665

666

667

668

669 **Figures**



670

671 **Fig. 1 Strategy for HCMV epitope-specific T cell identification:** PBMCs from HCMV

672 seropositive subjects were stimulated with 2 µg/ml pools and plated on IFN-γ coated

673 fluorospot plates for 20 hours. The top 10 positive pools (indicated by * on bars) were

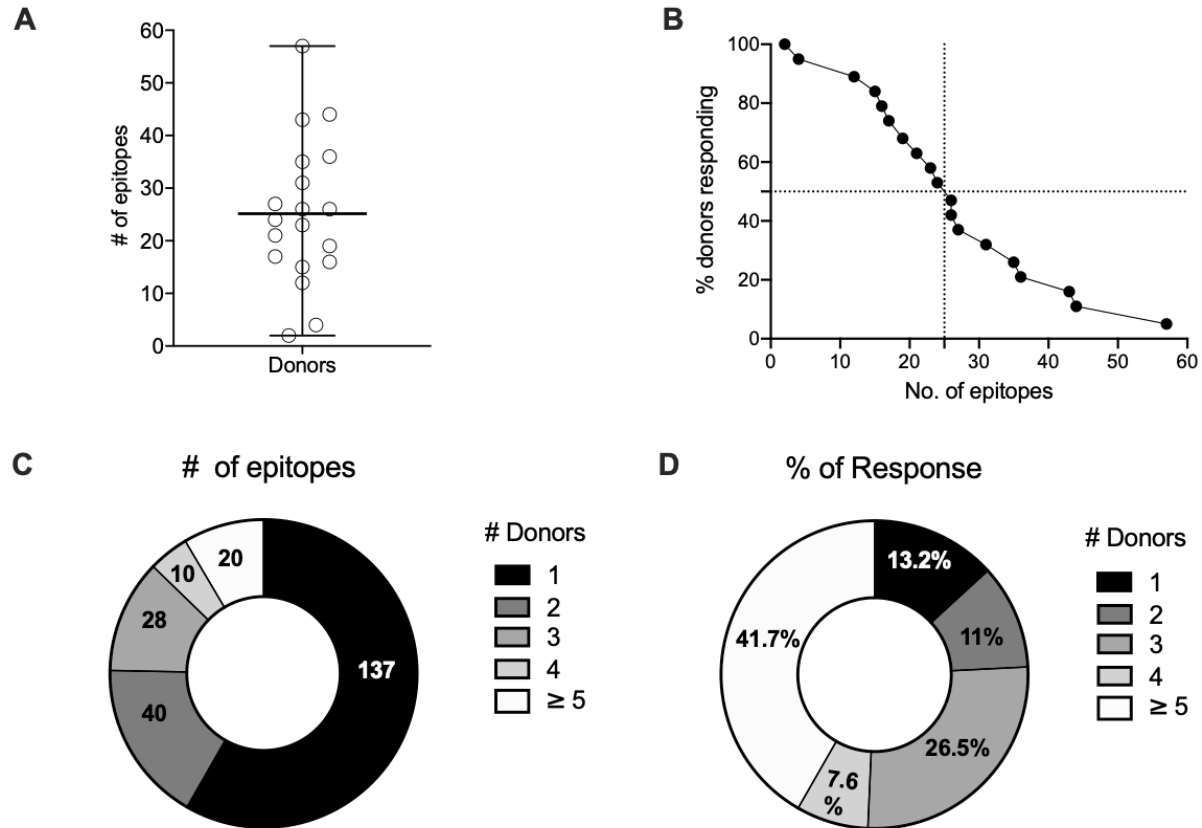
674 deconvoluted to identify individual epitopes. PBMC were stimulated with 10 µg/ml of each

675 individual peptide contained in the pool and reactivity was measured by IFN-γ fluorospot

676 assay. (A) SFC/10⁶ PBMC for one representative subject against the 89 peptide pools (B)

677 Deconvoluted pool representing individual peptides.

678



679
680

681 **Fig. 2 Breadth and dominance of HCMV T cell responses:** (A) The number of epitopes
682 recognized by each donor, mean \pm range. (B) Proportion of the 19 donors that responded
683 to the indicated number of epitopes. (C) Epitopes by number of responding donors. (D)
684 epitope % of total response by number of responding donors.

685

686

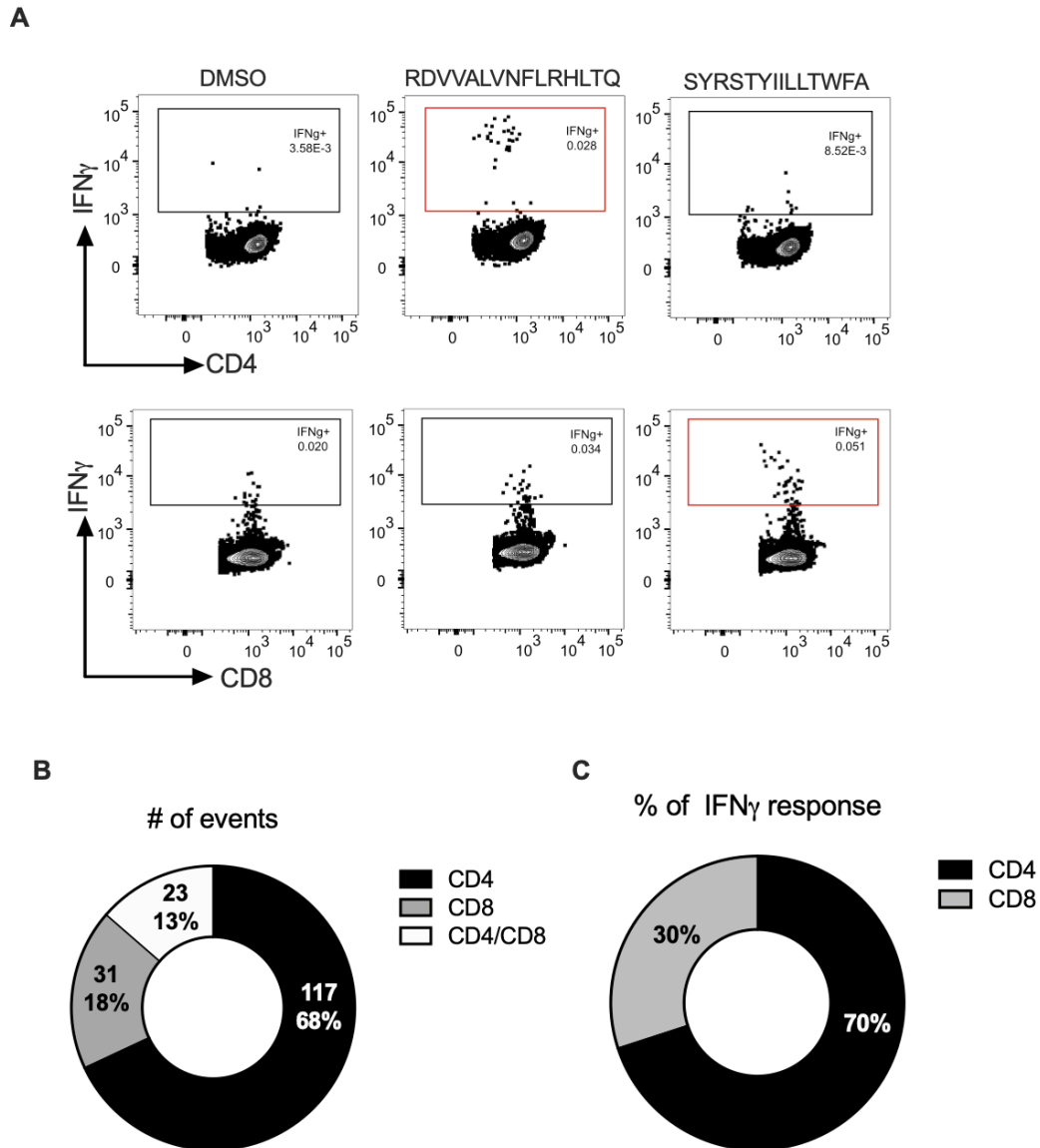
687

688

689

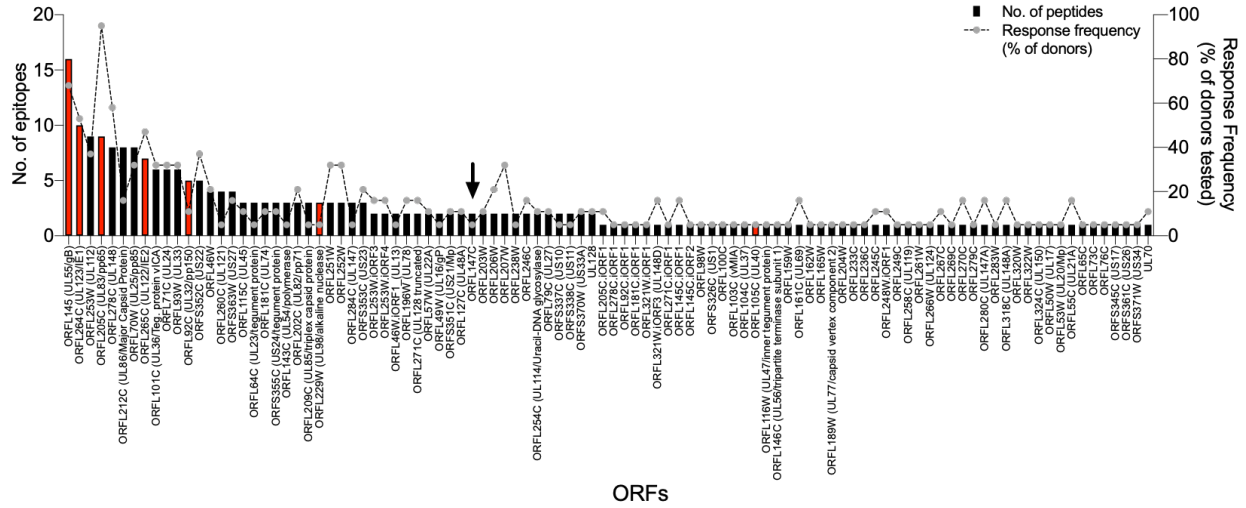
690

691



692

693 **Fig. 3 Phenotypic characterization of HCMV T cell responses:** (A) Representative
694 FACS plots for intracellular IFN- γ production by CD4+ and CD8+ T cells (gating axis in
695 red) upon stimulation with two of the scoring peptide epitopes that induced them (B and
696 C). The number of events and % response attributable to CD4+ and CD8+ T cell
697 responses of dominant epitopes (n=58) that demonstrated response frequency of 0.15
698 (15%) (i.e recognized by 3 or more donors).



699

700 **Fig. 4 T cell epitope distribution by ORF of origin: 235 epitopes mapped to 100**

701 ORFs. Left Y axis denotes the number of epitopes associated with each ORF (bars) and
 702 right Y axis denotes the response frequency associated with each ORF (dotted line).

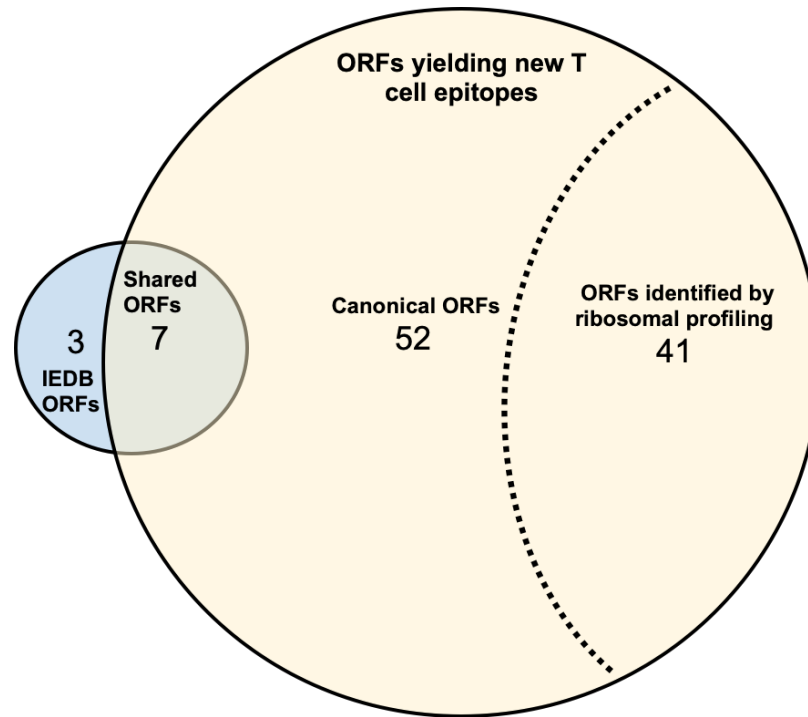
703 Seven canonical ORFs that were common in IEDB and the present screen are denoted
 704 in red. ORFL147C (arrow) is the first 'novel' ORF identified by rRNA profiling from left-
 705 to-right, and only induces responses in 2/19 individuals tested.

706

707

708

709



710

711 **Fig. 5 Overlap between IEDB-reported and new immunogenic ORFs identified in**
712 **this T cell epitope screen.** New epitopes were identified in all 100 ORFs, including 7
713 ORFs previously reported in the IEDB to be targets of T cell responses. Of the 93 ORFs
714 found to be new targets of T cells, 52 were canonical and 41 were 'novel' as identified
715 by recent ribosomal mRNA profiling studies.

716

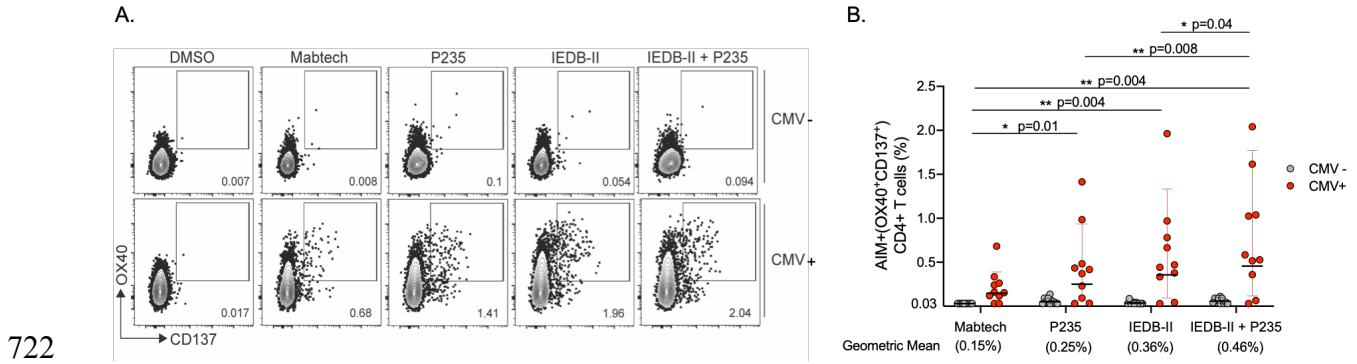
717

718

719

720

721



722

723 **Fig. 6 Epitope specific CD4+ T cell responses in HCMV+ and HCMV- subjects**

724 **detected with different peptide pools: (A) Representative FACS plots showing HCMV**

725 **specific CD4+ T cell reactivity against different peptide pools based on activation-induced**

726 **marker assays (OX40+ and CD137+ double expression). PBMCs from HCMV+ (red**

727 **circles) and CMV- donors (grey circles) were stimulated with 2 µg/ml of the Mabtech pool**

728 **or IEDB-II/P235 pools for 24 hrs. (B) Epitope-pool specific CD4+ T cells measured as**

729 **percentage of activation-induced marker assay positive (OX40+ CD137+) CD4+ T cells.**

730 **Each dot represents an individual subject. HCMV+ subjects demonstrated significantly**

731 **higher CD4+ T cell AIM responses than HCMV- subjects with all the different pools tested.**

732 **Mabtech HCMV+ vs HCMV- p=0.0007; P235 HCMV+ vs HCMV- p=0.0065; IEDB-II**

733 **CMV+ vs CMV- p=0.0009; P235/IEDB-II CMV+ vs CMV- p=0.004. Two-tailed Mann-**

734 **Whitney test. Comparisons across different pool formulations within the CMV+ were**

735 **made using the Wilcoxon matched-pairs signed ranked test, Two-tailed p values are**

736 **shown in the Figure; Geometric mean with geometric standard deviation.**

737

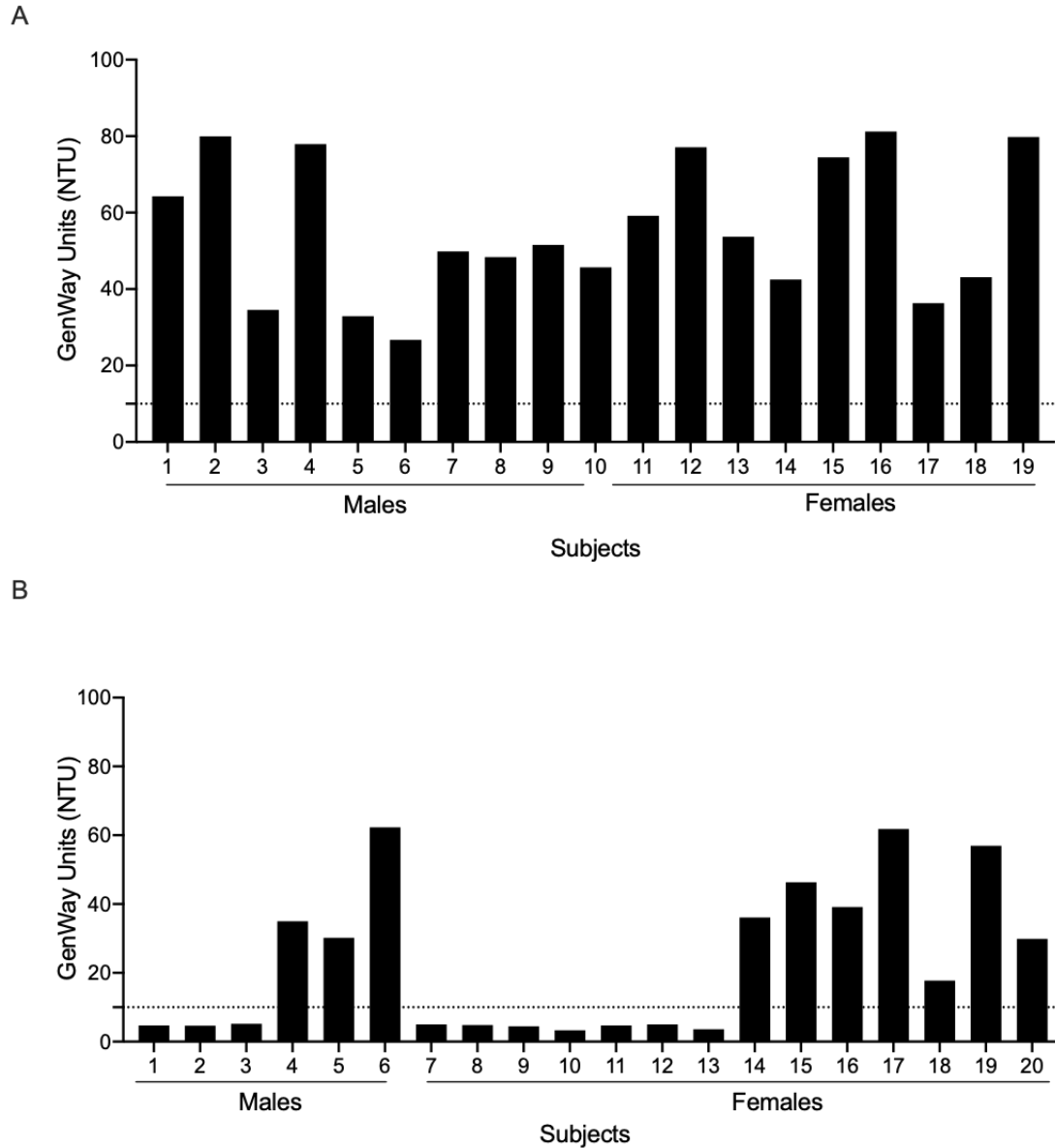
738

739

740

741 **Supplementary figures**

742



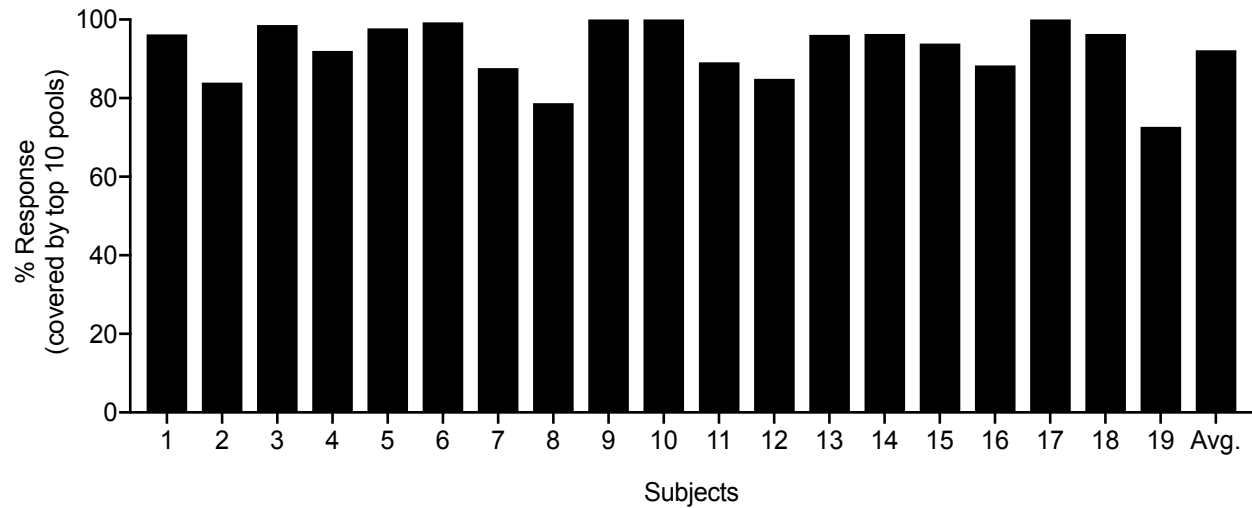
743

744 **Fig. S1 Confirmation of HCMV seropositivity in donors.** (A) IgG levels in plasma of

745 subjects in the screening cohort (n=10 males, n=9 females) and (B) the validation

746 cohort (n=13 males, n=26 females) determined by ELISA. Dotted line represents the cut

747 off for positivity (10 NTU).



748

749 **Fig. S2 The total T cell response captured by the top 10 epitope pools in each**
750 **subject.** The response magnitude of the top 10 pools as a percentage of the total
751 response magnitude observed from all positive pools. On average, the top 10 pools
752 accounted for ~90 % of each subject's total response.

753

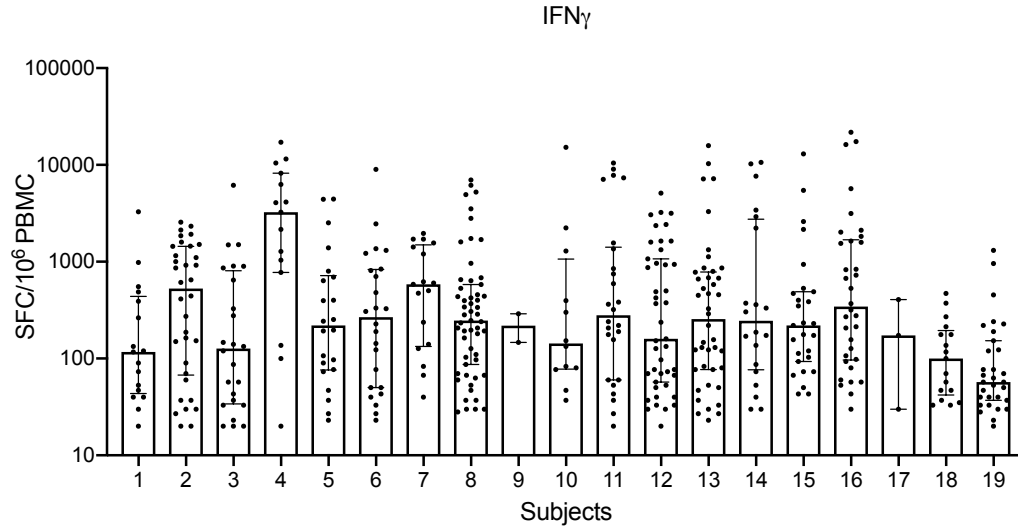
754

755

756

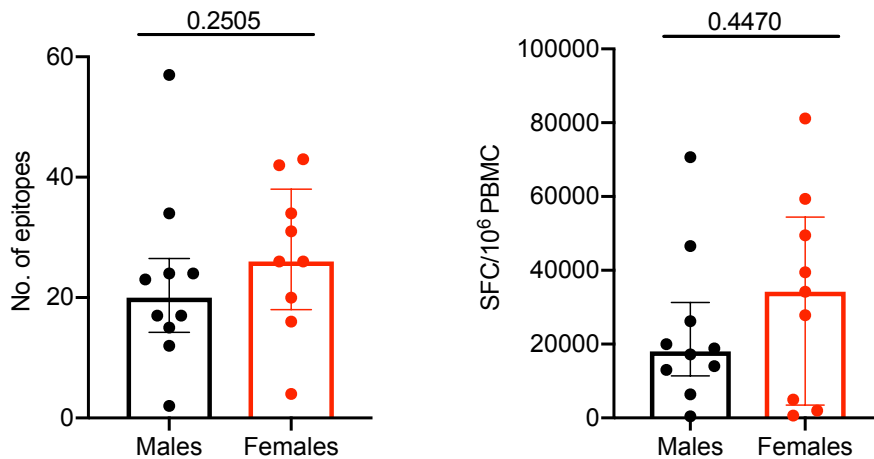
757

758



759

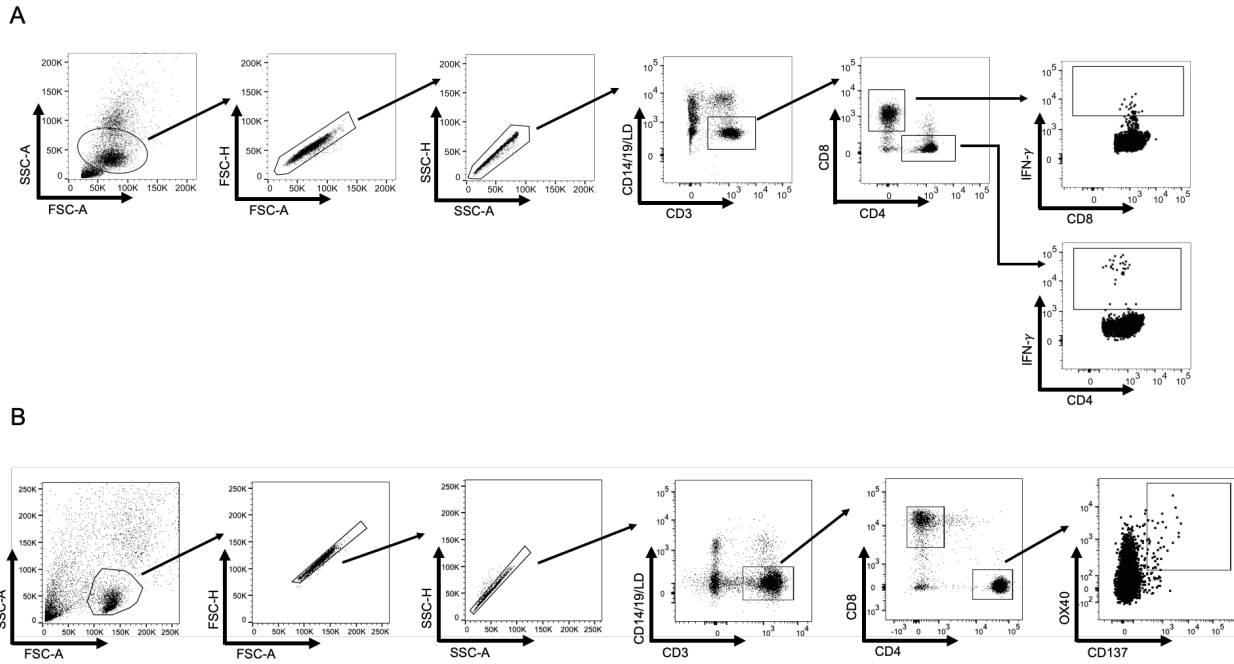
760 **Fig. S3 Response magnitude of each epitope identified in HCMV seropositive**
761 **individuals:** Each dot represents an epitope. Y axis represents the response magnitude
762 of individual epitopes. X axis represents each subject. Median \pm interquartile range is
763 shown.



764

765 **Fig. S4 Frequency and magnitude of response in males and females:** Each dot
766 represents a donor. Black dot/bar represents males and red dot/bar represents females.
767 Median with interquartile range is displayed. Two-tailed Mann-Whitney test.

768



769

770 **Fig. S5 Gating strategy adopted in IFN- γ Fluorospot and AIM assay:** (A) Human
771 PBMCs isolated from HCMV+ subjects were stimulated with each scoring peptide to
772 identify HCMV-specific IFN- γ producing CD4+ and CD8+ T cells. (B) Human PBMCs
773 isolated from HCMV+ and HCMV- subjects were stimulated with each megapool
774 generated to identify HCMV-specific activation-induced marker assay positive (OX40+
775 CD137+) CD4+ T cells.

776

777

778

779

780

781

782

783 **Table 1: Details of HCMV specific 235 epitopes identified in the screen.**

S. No	Peptide sequence	Peptide length	ORF(s)	No. of subjects responding	Magnitude of response
1	NGIRWQYQELQYLVE	15	ORFL46W, ORFL46W.iORF1_(UL13), ORFL46W.iORF2	2	222
2	RYNALTVRSRDSLLL	15	ORFL46W, ORFL46W.iORF1_(UL13), ORFL46W.iORF2	2	204
3	RVRTWVFVQRTTLWRR	15	ORFL46W, ORFL46W.iORF1_(UL13), ORFL46W.iORF2	1	60
4	GLWVSSYLVRPMTI	15	ORFL46W, ORFL46W.iORF1_(UL13), ORFL46W.iORF2	2	890
5	QGATYQLSIVRQAMQ	15	ORFL46W.iORF1_(UL13)	1	650
6	GAGLRQLRQQLTVRW	15	ORFL46W.iORF2	1	20
7	MRTVPVTKLYTSRMV	15	ORFL49W_(UL16)(UL16P?)	1	5677
8	AITLFFLLALRIPQ	15	ORFL49W.iORF1	1	207
9	ALFTHFVGRPRHCRL	15	ORFL50W_(UL17)	1	57
10	MLGIRAMLVMLDYYW	15	ORFL53W_(UL20)	1	350
11	PSVRMDFRARRPLRR	15	ORFL55C_(UL21A)	3	1740
12	ARRLWILSLLAVTLT	15	ORFL57W_(UL22A), ORFL57W.iORF1	1	100
13	LLAVTLTVALAAPSQ	15	ORFL57W_(UL22A)	2	6420
14	KDRCLVIRRRWRLVR	15	ORFL64C_(UL23)	1	60

15	FVAESITEFLNIGLR	15	ORFL64C_(UL23)	1	530
16	HENGIYYGTRSMRKL	15	ORFL64C_(UL23), ORFL64C.iORF1	1	827
17	FCRRFFFPDRPDFFL	15	ORFL65C	1	107
18	AEDSVFTSTRARSAT	15	ORFL70W_(UL25)	1	490
19	KFVLQDFDVQHLRRL	15	ORFL70W_(UL25)	1	40
20	IINYYYVAQKKARHM	15	ORFL70W_(UL25)	1	163
21	ALALHFLTSRKGVTD	15	ORFL70W_(UL25)	1	40
22	LMITHFQRTIRVLRC	15	ORFL70W_(UL25)	3	2421
23	DFLRVVRQQDAFICT	15	ORFL70W_(UL25)	2	174
24	ICVARLQAQPSSRHI	15	ORFL70W_(UL25)	1	37
25	GVSSVTLLKIFSQVP	15	ORFL70W_(UL25)	2	220
26	VLATLAAVRTRRRSV	15	ORFL71C, ORFL71C.iORF1 (UL24)	2	340
27	EAYVRINAGQVLPVV	15	ORFL71C, ORFL71C.iORF1 (UL24)	1	1853
28	LHCMRYLTSSLVKRY	15	ORFL71C	1	150
29	KRYFRPLLRAWLGL	15	ORFL71C, ORFL71C.iORF1 (UL24)	4	1388
30	HLLRNIKTAFGMRVL	15	ORFL71C, ORFL71C.iORF1 (UL24)	2	1093
31	ARNLMEFARVGLRAV	15	ORFL71C.iORF1 (UL24)	1	1450
32	TGLVLLLLLLLVVRL	15	ORFL73C	1	1440
33	MLFRPTISNSIPRCR	15	ORFL76C	1	47
34	LRIIRLLRASIRHEY	15	ORFL79C_(UL27)	1	810
35	RAHIQKFERLHVRRF	15	ORFL79C_(UL27)	1	2523
36	SLQFIGLQRRDVAL	15	ORFL92C_(UL32)	1	83

37	RDVVALVNFLRHLTQ	15	ORFL92C_(UL32)	1	1710
38	RRTVLFNELMLWLGY	15	ORFL92C_(UL32)	1	180
39	VNAVNKLVTGRLIM	15	ORFL92C_(UL32)	1	340
40	KELRMCLSFDSNYCR	15	ORFL92C_(UL32)	1	127
41	GMKTVAFDLSSPQKS	15	ORFL92C.iORF1	1	193
42	NAIVLITQLLTNRVL	15	ORFL93W_(UL33)	1	37
43	STNFLTLTVLPFIVL	15	ORFL93W_(UL33)	1	63
44	VLPFIVLSNQWLLPA	15	ORFL93W_(UL33)	1	247
45	FATVALIAADRYRVL	15	ORFL93W_(UL33)	4	2310
46	SYRSTYIILLTWFA	15	ORFL93W_(UL33)	3	3990
47	LTLRRTIGTLSRLVP	15	ORFL93W_(UL33)	1	163
48	RRRMVSVTLFSPYSV	15	ORFL98W.iORF1, ORFL98W.iORF2	1	53
49	GRLMEVRQRNGRLRR	15	ORFL100C	1	30
50	WPERCFIQLRSRSAL	15	ORFL101C, ORFL101C.iORF1_(UL36)	3	313
51	GPGFMRYQLIVLIGQ	15	ORFL101C, ORFL101C.iORF1_(UL36)	1	3167
52	IQTMELMIRTVPRIT	15	ORFL101C, ORFL101C.iORF1_(UL36)	2	384
53	EFLVRQYVLVDTFGV	15	ORFL101C	2	104
54	RREAIVRLEKTPTCQ	15	ORFL101C, ORFL101C.iORF1_(UL36)	2	407
55	RRRFKVCVGRRHII	15	ORFL101C, ORFL101C.iORF1_(UL36)	1	63
56	RHRFLWQRRRRARLL	15	ORFL103C_(vMIA), ORFL104C_(UL37)	1	1000

57	GSFSSFYSQIARSLG	15	ORFL105C_(UL40)	1	363
58	FLKKMLLCALKGRAS	15	ORFL115C_(UL45), ORFL115C.iORF1	1	930
59	MPVQRLTVNVARCVF	15	ORFL115C_(UL45)	1	20
60	KFIFELYRLPRLSIA	15	ORFL115C_(UL45)	1	173
61	ASKIKMLETRVTLAL	15	ORFL116W_(UL47)	1	140
62	ATMLSKYTRMSSLFN	15	ORFL127C_(UL48A)	2	650
63	AFKLDLLRMIAVSRT	15	ORFL127C_(UL48A)	2	477
64	MLFFQRYAPAFVTGY	15	ORFL143C_(UL54)	1	57
65	DLKYILTRLEYLYKV	15	ORFL143C_(UL54)	1	30
66	DPSYVREHGVPPIHAD	15	ORFL143C_(UL54)	1	123
67	TDLIRFERNIVCTSM	15	ORFL145C_(UL55)	3	273
68	EGIMVVYKRNIVAHT	15	ORFL145C_(UL55)	1	637
69	HTFKVRVYQKVLTR	15	ORFL145C_(UL55)	1	423
70	YQKVLTFRRSYAYIH	15	ORFL145C_(UL55)	1	40
71	RRSYAYIHTTYLLGS	15	ORFL145C_(UL55)	5	2730
72	QLMPDDYSNTHSTRY	15	ORFL145C_(UL55)	5	45517
73	NLNCMVTITTARSKY	15	ORFL145C_(UL55)	2	4810
74	NADKFFIFPNYIVS	15	ORFL145C_(UL55)	4	5120
75	GLVFWQGIKQKSLV	15	ORFL145C_(UL55)	1	53
76	QLQFTYDTLRGYINR	15	ORFL145C_(UL55)	3	2236
77	LRGYINRALAQIAEA	15	ORFL145C_(UL55)	3	30337
78	KELSKINPSAISAI	15	ORFL145C_(UL55)	3	18557
79	AILSAINKPIAARF	15	ORFL145C_(UL55)	5	11264
80	ASCVTINQTSVKVLR	15	ORFL145C_(UL55)	5	8941
81	YLFKRMIDLSSISTV	15	ORFL145C_(UL55)	1	67
82	EQAYQMLLALARLDA	15	ORFL145C_(UL55)	4	21807

83	LLDRLRHRKNGYRHL	15	ORFL145C.iORF1	3	12544
84	QILWTDGLARRTRDR	15	ORFL145C.iORF2	1	37
85	RVGITIQQLNVYHQL	15	ORFL146C_(UL56)	1	963
86	TMRSVFEMQRIRHGA	15	ORFL147C	1	67
87	NIFLVGFYLLVPYLG	15	ORFL147C	1	1633
88	SLILVLLLIYRCC	15	ORFL159W	1	23
89	LSYMKYHHLHGLPVN	15	ORFL161C_(UL69)	3	879
90	VELCLGAGAGHVVVV	15	ORFL162W	1	383
91	RSSWRASCVEVPKPKP	15	ORFL165W	1	370
92	MQKYFSLDNFLHDYV	15	UL70	2	20343
93	QTIYFLGLTALLRY	15	ORFL181C_(UL74)	1	583
94	SFYLVNAMSRLFRV	15	ORFL181C_(UL74)	1	43
95	TMRKLRKQALVKEQ	15	ORFL181C_(UL74)	1	1563
96	TAVSEFMKNTHVLIR	15	ORFL181C.iORF1	1	747
97	WREDVLMDRVVRKRYL	15	ORFL189W_(UL77)	1	67
98	IKMWFLGAPMIAVL	15	ORFL196W_(UL78), ORFL196W.iORF1	3	2867
99	LFIIAFFSREPTKDL	15	ORFL196W_(UL78)	1	190
100	PKSFTLTRIHPYIV	15	ORFL202C_(UL82/pp71)	2	340
101	PEYIVQIQNAFETNQ	15	ORFL202C_(UL82/pp71)	4	527
102	GALTLVIPSWHVFAS	15	ORFL202C_(UL82/pp71)	2	190
103	CRSATSLVGNTNADV	15	ORFL203W	1	30
104	SSCAHTTCRSATSLV	15	ORFL203W	2	104
105	SWLGQMLRPVGLCTL	15	ORFL204W	1	43
106	QTGIHVRVSQPSLIL	15	ORFL205C_(UL83/pp65), ORFL205C.iORF1	11	33563
107	MSIYVYALPLKMLNI	15	ORFL205C_(UL83/pp65)	1	83

108	PLKMLNIPSINVHHY	15	ORFL205C_(UL83/pp65)	7	2920
109	ATKMQVIGDQYVKVY	15	ORFL205C_(UL83/pp65), ORFL205C.iORF1	2	4637
110	PKNMIIKPGKISHIM	15	ORFL205C_(UL83/pp65), ORFL205C.iORF1	5	538
111	PGKISHIMLDVAFTS	15	ORFL205C_(UL83/pp65), ORFL205C.iORF1	9	2386
112	MNGQQIFLEVQAIRE	15	ORFL205C_(UL83/pp65), ORFL205C.iORF1	6	3990
113	ELRQYDPVAALFFFD	15	ORFL205C_(UL83/pp65)	8	1106
114	GILARNLVPMVATVQ	15	ORFL205C_(UL83/pp65), ORFL205C.iORF1	11	34633
115	ALFFFDIDLLLQRGP	15	ORFL205C.iORF1	2	73
116	RVTGLVFSVFSVSL	15	ORFL206W	4	6380
117	LTWCVIADRQPRFSV	15	ORFL206W	3	240
118	RPKRRVVAPFRVAAA	15	ORFL207W	2	283
119	APFRVAAAGETPLGR	15	ORFL207W	6	1089
120	IPQRLHLIKHYQLGL	15	ORFL209C_(UL85)	1	437
121	IVPMPLALEINQRLL	15	ORFL209C_(UL85)	1	283
122	LASELTMTYVRKLAL	15	ORFL209C_(UL85)	1	37
123	HSILADFNSYKAHLT	15	ORFL212C_(UL86) Major Capsid Protein	1	27
124	FHELRTWEIMEHMRL	15	ORFL212C_(UL86) Major Capsid Protein	1	7343
125	PQLLFHYRNLVAVLR	15	ORFL212C_(UL86) Major Capsid Protein	1	60

126	RNLVAVLRLVTRISA	15	ORFL212C_(UL86) Major Capsid Protein	1	20
127	LFLAVQFVGEHVKVL	15	ORFL212C_(UL86) Major Capsid Protein	1	53
128	VRVQDLFRVFP MNVY	15	ORFL212C_(UL86) Major Capsid Protein	1	43
129	LGYN SKFYSPCAQYF	15	ORFL212C_(UL86) Major Capsid Protein	1	20
130	TQEALPILSTTTLAL	15	ORFL212C_(UL86) Major Capsid Protein	1	407
131	PFTVLR LSYAYRIFA	15	ORFL229W_(UL98)	1	33
132	AREFLLSHDAALFRA	15	ORFL229W_(UL98)	1	23
133	MLIQQYVLSQYYIKK	15	ORFL229W_(UL98)	1	20
134	RLGTAATQIQKQTLY	15	ORFL233C	1	30
135	KTQIFNKLF TNRISV	15	ORFL236C	1	1587
136	VRSLAVDAQHAAKRV	15	ORFL238W	1	53
137	LEERDEWVRSLAVDA	15	ORFL238W	1	47
138	AAITVVPVITQSRLA	15	ORFL245C	3	450
139	PWYPITQARTLELTP	15	ORFL246C	2	996
140	MSTKRSTVPWYPITQ	15	ORFL246C	3	477
141	LRVTFHRVKPTLQRE	15	ORFL248W.iORF1	2	830
142	SGRVILWTTLR LRCIL	15	ORFL249C	1	20
143	VVRKYWTF TNP NRIL	15	ORFL251W, ORFL252W, ORFL253W_(UL112), ORFL253W.iORF1, ORFL253W.iORF2	3	16876

144	TFDVRQFVFDNARLV	15	ORFL251W, ORFL252W, ORFL253W_(UL112), ORFL253W.iORF1, ORFL253W.iORF2	6	13736
145	VRGGIVFNKSVSSVV	15	ORFL251W, ORFL252W, ORFL253W_(UL112), ORFL253W.iORF1, ORFL253W.iORF2	5	2691
146	GNLQVTYVRHYLKNH	15	ORFL253W_(UL112), ORFL253W.iORF1, ORFL253W.iORF2	4	655
147	AVAFLNYSSSSSAVS	15	ORFL253W_(UL112), ORFL253W.iORF1, ORFL253W.iORF2	3	2151
148	AGLMMRMRRAPAE	15	ORFL253W_(UL112)	1	490
149	CDLPLVSSRLLPETS	15	ORFL253W_(UL112)	1	123
150	CEIKPYVNPVATA	15	ORFL253W_(UL112)	3	2583
151	DPLLRSLQVAGSGRR	15	ORFL253W_(UL112)	2	2067
152	LPLCSTARLRLAPRR	15	ORFL253W.iORF3	1	467
153	RATGNFRSTSLYAAV	15	ORFL253W.iORF3	3	4220
154	RCCTLRFRRCRARC	15	ORFL253W.iORF4	2	713
155	MSATRHRCCTLRFR	15	ORFL253W.iORF4	1	277
156	RVFCLSADWIRFLSL	15	ORFL254C_(UL114)	2	846
157	HLGWQTLNHHVIRRL	15	ORFL254C_(UL114)	1	127
158	TVVRLHVQIAGRSFT	15	ORFL258C_(UL119)	1	213
159	SCTHPYVSLVTPLT	15	ORFL260C_(UL121)	1	80
160	ISLVTPLTINATLRL	15	ORFL260C_(UL121)	1	317

161	CRVDADLGLLYAVCL	15	ORFL260C_(UL121)	1	673
162	VCLILSFSIVTAALW	15	ORFL260C_(UL121)	1	43
163	MFFLAIRDHDTAGGI	15	ORFL261W	1	1637
164	LQTMLRKEVNSQLSL	15	ORFL264C_(UL123) IE1, ORFL265C_(UL122) IE2	2	1606
165	LVKQIKVRVDMVRHR	15	ORFL264C_(UL123) IE1	3	580
166	RVDMVRHRIKEHMLK	15	ORFL264C_(UL123) IE1	2	250
167	LRRKMMYMCYRNIEF	15	ORFL264C_(UL123) IE1	6	1697
168	CSPDEIMSYAQKIFK	15	ORFL264C_(UL123) IE1	2	194
169	EERDKVLTHIDHIFM	15	ORFL264C_(UL123) IE1	2	107
170	VLCCYVLEETSVMLA	15	ORFL264C_(UL123) IE1	1	93
171	ITKPEVISVMKRRIE	15	ORFL264C_(UL123) IE1	1	1600
172	FAQYILGADPLRVCS	15	ORFL264C_(UL123) IE1	1	30
173	EAIWAYTLATAGASS	15	ORFL264C_(UL123) IE1	3	19333
174	TTRPFKVIKPPVPP	15	ORFL265C_(UL122) IE2	2	424
175	NKGIQIIYTRNHEVK	15	ORFL265C_(UL122) IE2, ORFL265C.iORF1, ORFL265C.iORF2, ORFL265C.iORF3,	7	3794
176	LGSMCNLALSTPFLM	15	ORFL265C_(UL122) IE2, ORFL265C.iORF1, ORFL265C.iORF2	4	668
177	STPFLMEHTMPVTHP	15	ORFL265C_(UL122) IE2, ORFL265C.iORF1, ORFL265C.iORF2, ORFL265C.iORF3,	4	740
178	YRNMIIHAATPVDLL	15	ORFL265C.iORF3	2	130

179	VMVRIFSTNQGGFML	15	ORFL265C.iORF3	2	3560
180	VVVGIVLCLSLASTV	15	ORFL266W_(UL124)	1	1417
181	SPVAAELPHSPAPM	15	ORFL267C	2	166
182	SYLAVHLRISHRYH	15	ORFL269C	1	290
183	IAITMVMRFWQYING	15	ORFL270C	3	163
184	TALWLLLGHSRVPV	15	UL128	1	177
185	AEIRGIVTTMTHSLT	15	ORFL271C_(UL128_truncated)	1	1713
186	NPLYLEADGRIRCGK	15	ORFL271C_(UL128_truncated), UL128	2	884
187	LHRRAAVSGRRSLLQ	15	ORFL271C.iORF1	1	87
188	MLRLLFTLVLLALYG	15	ORFL278C_(UL148)	3	4593
189	HVRLLSYRGDPLVFK	15	ORFL278C_(UL148)	3	333
190	VVRFALYLETLRIV	15	ORFL278C_(UL148)	2	123
191	FYMNWTLRRSQTHYL	15	ORFL278C_(UL148)	6	16883
192	QVEILKPRGVRHRAI	15	ORFL278C_(UL148)	9	3468
193	FCVYRYNARLTRGYV	15	ORFL278C_(UL148)	3	700
194	TRGYVRYTLSPKARL	15	ORFL278C_(UL148)	7	9337
195	SLDRFIVQYLNTLLI	15	ORFL278C_(UL148)	8	23368
196	PTWSTTVNAHNSFLH	15	ORFL278C.iORF1	1	47
197	DRLSTLAATMCMFDY	15	ORFL279C	1	53
198	LFYRAVALGTL SALV	15	ORFL280C_(UL147A)	3	2633
199	SSIFTSTHRGIVIVAP	15	ORFL283W	1	27
200	LSVRYLSLTAYMLLA	15	ORFL284C_(UL147)	1	1200
201	TAYKAFLWKYAKKLN	15	ORFL284C_(UL147)	1	503
202	WKYAKKLNHYFRLR	15	ORFL284C_(UL147)	1	237
203	VYLWYVRRQLVAFCL	15	ORFL318C_(UL148A)	3	253
204	FPSARDIPKQLPEQP	15	ORFL320W	1	27

205	VVAYVILERLWLAAR	15	ORFL321W.iORF1	1	23
206	IRRWWISVAIVIFIG	15	ORFL321W.iORF2, ORFL321W.iORF3_(UL148D)	3	10480
207	RWQFAVCAASKTATR	15	ORFL322W	1	50
208	PQRLLLTALAIWQRT	15	ORFL324C_(UL150)	1	983
209	PWWRRLRVKRPKFPS	15	ORFS326C, ORFS326C.iORF1_(US1)	1	240
210	LWYLGDYGAILKIYF	15	ORFS337C_(US10)	1	40
211	LFCGACVITRSLLLI	15	ORFS337C_(US10)	1	487
212	MNLVMLLALWAPVA	15	ORFS338C_(US11)	1	553
213	VSEYRVEYSEARCVL	15	ORFS338C_(US11)	1	263
214	MLVVTVFDTTRLF EI	15	ORFS345C_(US17)	1	840
215	VCAFCWLVLPHRLEQ	15	ORFS351C_(US21)	1	1960
216	VSVLYFMPSEPGSAH	15	ORFS351C.iORF2	1	177
217	VFQKTL SMLQGLYLR	15	ORFS352C_(US22)	2	327
218	GLYLRQYDPPALRTY	15	ORFS352C_(US22)	2	633
219	WFLVMREQAAIPQIY	15	ORFS352C_(US22)	4	1100
220	QIYARSLAADYLCCD	15	ORFS352C_(US22)	1	33
221	DFRDLLNFIRQRLCC	15	ORFS352C_(US22)	1	30
222	PSQEILLLCARHLDE	15	ORFS353C_(US23)	3	110
223	TDCWPFEVAPAARLA	15	ORFS353C_(US23)	2	1637
224	LFRAGLMKVYVRRRY	15	ORFS353C_(US23)	1	870
225	VVFMGRFSRVYAYDT	15	ORFS355C_(US24)	1	70
226	EKYMVLVSHNLDELA	15	ORFS355C_(US24)	1	20
227	PRLHCLVTTRSSTRE	15	ORFS355C.iORF1	1	1277
228	LRYKWLIRKDRFIVR	15	ORFS361C_(US26)	1	3247
229	TNIMLQVSNVTNHTL	15	ORFS363W_(US27)	1	58

230	IVVGLPFFLEYAKHH	15	ORFS363W_(US27)	2	1250
231	YNRMVRFIINYVGKW	15	ORFS363W_(US27)	1	23
232	ITFCLYVGQFLAYVR	15	ORFS363W_(US27)	1	27
233	HDPLGLTRFIMRQLM	15	ORFS370W_(US33A)	1	473
234	FIMRQLMMYPLVLPF	15	ORFS370W_(US33A)	1	530
235	GLVYRELHDFYGYLQ	15	ORFS371W_(US34)	1	963

784

785

786

787

788

789

790

791

792

793

794

795

796

797

798

799

800

801

802 **Table 2: Details of HCMV specific class II epitopes from IEDB**

S. No	Peptide sequence	Peptide length	ORF	Antigen Name from IEDB
1	HINSHSQCYSYSRVIA	17	ORFL145C_(UL55)	glycoprotein B
2	SRVIAGTVFVAYHRD	15	ORFL145C_(UL55)	glycoprotein B
3	CMVTITTARSKYPYH	15	ORFL145C_(UL55)	glycoprotein B
4	VFETTGGLVFWQGI	15	ORFL145C_(UL55)	glycoprotein B
5	MQLIPDDYSNTHSTRYVTVK	20	ORFL145C_(UL55)	glycoprotein B
6	LPLKMLNIPSINVH	14	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein
7	PQYSEHPTFTSQYRIQ	16	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein
8	FTSQYRIQGKLEYRHT	16	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein
9	PPWQAGILARNLVPMV	16	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein
10	KYQEFFWDANDIYRIF	16	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein
11	GPISGHVLKAVFSRG	15	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein
12	LLQTGIHVRVSQPSL	15	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein
13	IYVYALPLKMLNIPS	15	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein

14	LPLKMLNIPSINVHH	15	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein
15	KDVALRHVVCAHELV	15	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein
16	RHVVCAHELVCSMEN	15	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein
17	CSMENTRATKMQVIG	15	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein
18	TRATKMQVIGDQYVK	15	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein
19	MQVIGDQYVKVYLES	15	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein
20	VYLESFCEDVPSGKL	15	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein
21	FCEDVPSGKLFMHVT	15	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein
22	LGSDVEEDLTMTRNP	15	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein
23	EEDLTMTRNPQPFMR	15	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein
24	QPFMRPHERNGFTVL	15	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein
25	KISHIMLDVAFTSHE	15	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein
26	MLDVAFTSHEHFGLL	15	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein

27	FTSHEHFGLLCPKSI	15	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein
28	PQYSEHPTFTSQYRI	15	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein
29	SQYRIQGGKLEYRHTW	15	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein
30	YRHTWDRHDEGAAQG	15	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein
31	IHNPAVFTWPPWQAG	15	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein
32	PWQAGILARNLVPMV	15	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein
33	ATVQQGNLKYQEFFF	15	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein
34	QNLKYQEFFFWDANDI	15	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein
35	QEFFFWDANDIYRIFA	15	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein
36	ELEGVWQPAAQPKRR	15	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein
37	IFLEVQAI RETVELR	15	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein
38	PPWQAGILARNLVPM	15	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein
39	DVPSGKLFMHVTLGS	15	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein

40	KLFMHVTLGSDVEED	15	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein
41	DVEEDLTMTRNPQPF	15	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein
42	VAFTSHEHFGLLCPK	15	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein
43	SEHPTFTSQYRIQ GK	15	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein
44	LEYRHTWDRHDEGAA	15	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein
45	PLKMLNIPSINVHHY	15	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein
46	KVYLESFCEDVPSGK	15	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein
47	TLGSDVEEDLTMTRN	15	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein
48	ASTSAGRKRKSASSA	15	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein
49	ACTSGVMTRGRLKAE	15	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein
50	AGILARNLVPMVATV	15	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein
51	EPDVYYTSAFVFPTK	15	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein
52	QVIGDQYVKVYLESF	15	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein

53	FFWDANDIYRIFAEI	15	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein
54	LVSQYTPDSTPCHRG	15	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein
55	SHIMLDVAFTSHEH	14	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein
56	DESDNEIHNPVFTW	16	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein
57	SQYTPDSTPCHRG	13	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein
58	KPGKISHIMLDVA	13	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein
59	PTFTSQYRIQGKL	13	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein
60	DTPVLPHETRLQLTGIVRV	20	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein
61	INVHHYPSAAERKHRHLPVA	20	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein
62	LLQRGPQYSEHPTFT	15	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein
63	ALFFFDIDLLLQRGPQYSE	19	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein
64	DQYVKVYLESFCEDVPSGKL	20	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein
65	MTRNPQPFMRPHERNGFTV L	20	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein

66	MISVLGPISGHV LKAVFSRG	20	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein
67	ASGKQMWQARLTVSGLAWT R	20	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein
68	LPLKMLNIPSINVHHYPSAA	20	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein
69	PHETRLQLTGIVRV SQPSL	20	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein
70	IYVYALPLKMLNIPSINVHH	20	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein
71	QYDPVAALFFFDIDL LQRG	20	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein
72	RQYDPVAALFFFDIDL	16	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein
73	HETRLQLTGIVRV S	15	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein
74	VYALPLKMLNIPSIN	15	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein
75	VALRHVVCAHELVCS	15	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein
76	HIMLDVAFTSHEHFG	15	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein
77	FTSQYRIQ GKLEYRH	15	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein
78	YRIQ GKLEYRHTWDR	15	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein

79	ARNLVPMVATVQGQN	15	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein
80	ANDIYRIFAELEGVW	15	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein
81	TRQQNQWKEPDVYYT	15	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein
82	TERKTPRVTTGGGAMA	15	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein
83	NLKYQEFFFWDANDIY	15	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein
84	TPRVTTGGGAMAGAST	15	ORFL205C_(UL83/pp65)	65 kDa lower matrix phosphoprotein
85	DQYVKVYLESFCEDV	15	ORFL205C_(UL83/pp65)	HCMVUL83
86	GKISHIMLDVAFTSH	15	ORFL205C_(UL83/pp65)	HCMVUL83
87	EHPTFTSQYRIQGKL	15	ORFL205C_(UL83/pp65)	HCMVUL83
88	GQNLKYQEFFFWDAND	15	ORFL205C_(UL83/pp65)	HCMVUL83
89	KYQEFFFWDANDIYRI	15	ORFL205C_(UL83/pp65)	HCMVUL83
90	IIPGKISHIMLDVA	15	ORFL205C_(UL83/pp65)	HCMVUL83
91	TRATKMQVIGDQYVKVYLES	20	ORFL205C_(UL83/pp65)	HCMVUL83
92	KLFMHVTLGSDVEEDLTMTR	20	ORFL205C_(UL83/pp65)	HCMVUL83
93	KPGKISHIMLDVAFTSHEHF	20	ORFL205C_(UL83/pp65)	HCMVUL83
94	LPVADAVIHASGKQMWQARL	20	ORFL205C_(UL83/pp65)	HCMVUL83
95	GSDSDEELVTTERKTPRVTTG	20	ORFL205C_(UL83/pp65)	HCMVUL83
96	RHRQDALPGPCIASTPKKHR	20	ORFL205C_(UL83/pp65)	HCMVUL83
97	YQEFFFWDANDIYR	13	ORFL205C_(UL83/pp65)	HCMVUL83
98	LAWTRQQNQWKEPDV	15	ORFL205C_(UL83/pp65)	HCMVUL83
99	YQEFFFWDANDIYRIF	15	ORFL205C_(UL83/pp65)	HCMVUL83

100	EFFWDANDIYRIF	13	ORFL205C_(UL83/pp65)	HCMVUL83
101	VEEDLTMTRNPQPFM	15	ORFL205C_(UL83/pp65)	HCMVUL83
102	KPGKISHIMLDVAFTSH	17	ORFL205C_(UL83/pp65)	HCMVUL83
103	TSQYRIQGKLEYRHT	15	ORFL205C_(UL83/pp65)	HCMVUL83
104	MSIYVYALPLKMLNI	15	ORFL205C_(UL83/pp65)	HCMVUL83
105	VYYTSAFVFPTKDVA	15	ORFL205C_(UL83/pp65)	HCMVUL83
106	LRQYDPVAALFFFDI	15	ORFL205C_(UL83/pp65)	HCMVUL83
107	GPQYSEHPTFTSQYRI	16	ORFL205C_(UL83/pp65)	HCMVUL83
108	HPTFTSQYRIQGKLE	15	ORFL205C_(UL83/pp65)	HCMVUL83
109	TRLLQTGIHVRVSQP	15	ORFL205C_(UL83/pp65)	HCMVUL83
110	RNGFTVLCPKNMIK	15	ORFL205C_(UL83/pp65)	HCMVUL83
111	PISGHVLKAVFSRGD	15	ORFL205C_(UL83/pp65)	HCMVUL83
112	GIHVRVSQPSLILVS	15	ORFL205C_(UL83/pp65)	HCMVUL83
113	IHASGKQMWQARLTV	15	ORFL205C_(UL83/pp65)	HCMVUL83
114	GKQMWQARLTVSGLA	15	ORFL205C_(UL83/pp65)	HCMVUL83
115	ENTRATKMQVIGDQY	15	ORFL205C_(UL83/pp65)	HCMVUL83
116	ATKMVIGDQYVKVY	15	ORFL205C_(UL83/pp65)	HCMVUL83
117	RIPHERNGFTVLCPKN	15	ORFL205C_(UL83/pp65)	HCMVUL83
118	AQGDDVWTSGSDSD	15	ORFL205C_(UL83/pp65)	HCMVUL83
119	SSATACTSGVMTRGR	15	ORFL205C_(UL83/pp65)	HCMVUL83
120	YRIFAELEGVWQPAA	15	ORFL205C_(UL83/pp65)	HCMVUL83
121	AELEGVWQPAAQPKR	15	ORFL205C_(UL83/pp65)	HCMVUL83
122	AVFSRGDTPVLPHEH	15	ORFL205C_(UL83/pp65)	phosphorylated matrix protein (pp65)
123	ALPLKMLNIPSINVH	15	ORFL205C_(UL83/pp65)	pp65
124	HVLKAVFSRGDTPVL	15	ORFL205C_(UL83/pp65)	pp65
125	AHELVCSMENTRATKMVIG	20	ORFL205C_(UL83/pp65)	tegument protein pp65

126	FCEDVPSGKLFMHVTLGSDV	20	ORFL205C_(UL83/pp65)	tegument protein pp65
127	TLGSDVEEDLTMTRNPQPF	19	ORFL205C_(UL83/pp65)	tegument protein pp65
128	LLQTGIHVRVVSQPSLILV	18	ORFL205C_(UL83/pp65)	tegument protein pp65
129	SICPSQEPMSIYVYA	15	ORFL205C_(UL83/pp65)	tegument protein pp65
130	SQEPMSIYVYALPLK	15	ORFL205C_(UL83/pp65)	tegument protein pp65
131	LNIPSINVHHYPSAA	15	ORFL205C_(UL83/pp65)	tegument protein pp65
132	HDVSKGDDNKLGGALQAKA	19	ORFL264C_(UL123) IE1	55 kDa immediate-early protein 1
133	ALQAKARDKKDELRRKMMY	19	ORFL264C_(UL123) IE1	55 kDa immediate-early protein 1
134	KEHMLKKYTQTEEFK	15	ORFL264C_(UL123) IE1	55 kDa immediate-early protein 1
135	QTEEFKFTGAFNMMGGCLQN	19	ORFL264C_(UL123) IE1	55 kDa immediate-early protein 1
136	MGGCLQNALDILDKVHEPFE	20	ORFL264C_(UL123) IE1	55 kDa immediate-early protein 1
137	AIVAYTLATAGVSSDSLIV	19	ORFL264C_(UL123) IE1	55 kDa immediate-early protein 1
138	TMQSMYENYIVPEDKREMW	19	ORFL264C_(UL123) IE1	55 kDa immediate-early protein 1
139	RRKMMYMCYRNIEFFTKNS	19	ORFL264C_(UL123) IE1	55 kDa immediate-early protein 1
140	FFTKNSAFPKTNGCSQAM	19	ORFL264C_(UL123) IE1	55 kDa immediate-early protein 1
141	CVETMCNEYKVTSDACMMT	19	ORFL264C_(UL123) IE1	55 kDa immediate-early protein 1

142	DACMMTMYGGASLLSEFCR	19	ORFL264C_(UL123) IE1	55 kDa immediate-early protein 1
143	NYIVPEDKREMWMACIKELH	20	ORFL264C_(UL123) IE1	55 kDa immediate-early protein 1
144	VRHRIKEHMLKKYTQTEEF	20	ORFL264C_(UL123) IE1	55 kDa immediate-early protein 1
145	VRVDMVRHRIKEHML	15	ORFL264C_(UL123) IE1	55 kDa immediate-early protein 1
146	VKQIKVRVDMVRHRI	15	ORFL264C_(UL123) IE1	55 kDa immediate-early protein 1
147	VRHRIKEHMLKKYTQ	15	ORFL264C_(UL123) IE1	55 kDa immediate-early protein 1
148	EQSDEEEEEGAQEER	15	ORFL264C_(UL123) IE1	55 kDa immediate-early protein 1
149	VKSEPVSEIEEVAPE	15	ORFL264C_(UL123) IE1	55 kDa immediate-early protein 1
150	PVSEIEEVAPEEEEEED	15	ORFL264C_(UL123) IE1	55 kDa immediate-early protein 1
151	LQNALDILDKVHEPF	15	ORFL264C_(UL123) IE1	55 kDa immediate-early protein 1
152	EDKREMWMACIKELH	15	ORFL264C_(UL123) IE1	55 kDa immediate-early protein 1
153	THIDHIFMDILTTCV	15	ORFL264C_(UL123) IE1	55 kDa immediate-early protein 1
154	VLEETSVMLAKRPLI	15	ORFL264C_(UL123) IE1	55 kDa immediate-early protein 1

155	TKPEVISVMKRRIEE	15	ORFL264C_(UL123) IE1	55 kDa immediate-early protein 1
156	RRIEEICMKVFAQYI	15	ORFL264C_(UL123) IE1	55 kDa immediate-early protein 1
157	NIEFFTKNSAFP KTT	15	ORFL264C_(UL123) IE1	regulatory protein IE1
158	LTHIDHIFMDILTTCVETM	19	ORFL264C_(UL123) IE1	regulatory protein IE1
159	AIVAYTLATAGASSDSL V	19	ORFL264C_(UL123) IE1	UL123; IE1
160	VRVDMVRHRIKEHMLK KYTQ	20	ORFL264C_(UL123) IE1	UL123; IE1
161	DKREMWMACIKELH	14	ORFL264C_(UL123) IE1	UL123; IE1
162	QSMYENYIVPEDKREM WMA C	20	ORFL264C_(UL123) IE1	UL123; IE1
163	TRRGRVKIDEVSRMF	15	ORFL265C_(UL122) IE2	45 kDa immediate-early protein 2
164	GDILAQAVNHAGIDS	15	ORFL265C_(UL122) IE2	45 kDa immediate-early protein 2
165	KTTRPFKVIKPPVP	15	ORFL265C_(UL122) IE2	45 kDa immediate-early protein 2
166	FKVIKPPVPPAPIM	15	ORFL265C_(UL122) IE2	45 kDa immediate-early protein 2
167	PEPDFTIQYRNKIID	15	ORFL265C_(UL122) IE2	45 kDa immediate-early protein 2
168	PFTIPSMHQVLDEAI	15	ORFL265C_(UL122) IE2	45 kDa immediate-early protein 2
169	LMQKFQVMVRIFS	15	ORFL265C_(UL122) IE2	45 kDa immediate-early protein 2
170	VRIFSTNQGGFMLPI	15	ORFL265C_(UL122) IE2	45 kDa immediate-early protein 2

171	PEDLDTLSLAIEAAI	15	ORFL265C_(UL122) IE2	45 kDa immediate-early protein 2
172	TLSLAIEAAIQDLRN	15	ORFL265C_(UL122) IE2	45 kDa immediate-early protein 2
173	SMHQVLDEAIKACKT	15	ORFL265C_(UL122) IE2	45 kDa immediate-early protein 2
174	KGIQIYTRNHEVKS	15	ORFL265C_(UL122) IE2	45 kDa immediate-early protein 2
175	ALSTPFLMEHTMPVT	15	ORFL265C_(UL122) IE2	45 kDa immediate-early protein 2
176	FLMEHTMPVTHPPEV	15	ORFL265C_(UL122) IE2	45 kDa immediate-early protein 2
177	PYAVAFQPLLAYAY	14	UL57	single-stranded DNA-binding protein
178	KTQLNRHSYLKDSDFLDAA	19	UL75	envelope glycoprotein H
179	RQTEKHELLVLVKKKAQLNRH	20	UL75	Glycoprotein H precursor
180	LDPHAFHLLLNTYGRPIR	18	UL75	Glycoprotein H precursor
181	KAQLNRHSYLKDSDFLDAA	19	UL75	Glycoprotein H precursor
182	DVLKSGRCQMLDRRTVEMA	19	UL75	Glycoprotein H precursor
183	LDKAFHLLLNTYGRPIR	17	UL75	Glycoprotein H precursor
184	KDQLNRHSYLKDPDFLDAA	19	UL75	Glycoprotein H precursor
185	SYLKDSDFLDAAL	13	UL75	HCMVUL75
186	RRIPHFYRVRREVPRTVNE	19	UL86	Major capsid protein
187	MDVNYFKIPNNPRGRASC	19	UL86	Major capsid protein

803

804

805

806 **Table 3: Demographic characteristics of HCMV (+/-) subjects analyzed in**
 807 **screening and validation studies.**

808

Characteristics	<i>Screening cohort</i>	<i>Validation cohort</i>	
	HCMV+	HCMV+	HCMV- 810
Total participants enrolled, n	19	10	10 811
Males/females	10/9	3/7	3/7 812
Median age (range)	65 (28-80)	35.5 (22-55)	28.5 (19-42) 813
Caucasian, % (n)	68 (13)	40 (4)	40 (4) 814
Asian, % (n)	16 (3)	10 (1)	20 (2) 815
African American, % (n)	5 (1)	10 (1)	10 (1) 816 817
More than one race, % (n)	0 (0)	30 (3)	30 (3) 818
Unknown, % (n)	10 (2)	10 (1)	0 (0)